MEPS

Methodology Report #31

Sample Design of the 2017 Medical Expenditure Panel Survey Insurance Component

Abstract

The primary purpose of this report is to describe the sample design, sample allocation, and sample selection process for the 2017 MEPS Insurance Component (MEPS-IC). This information is important for researchers using the data who wish to understand the details of its sampling design. Following a brief overview, both the private-sector and public (State and local governments) sector designs are described. The details presented in this report apply specifically to the 2017 data year, however the appendices include a history of sample allocation changes to the MEPS-IC.

<p align="center">* * *</p>

The estimates in this report are based on the most recent data available at the time the report was written. However, selected elements of MEPS data may be revised on the basis of additional analyses, which could result in slightly different estimates from those shown here. Please check the MEPS Web site for the most current file releases.

Background

The Medical Expenditure Panel Survey (MEPS)

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ), formerly the Agency for Health Care Policy and Research, and the National Center for Health Statistics (NCHS).

MEPS comprises three component surveys: the Household Component (HC), the Medical Provider Component (MPC), and the Insurance Component (IC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES) was conducted in 1977 and, the National Medical Expenditure Survey (NMES) in 1987. Beginning in 1996, MEPS continued this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features include linkage with the National Health Interview Survey (NHIS), from which the sample for the MEPS-HC is drawn, and enhanced longitudinal data collection for core survey components. The MEPS-HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

Household Component

The MEPS-HC, a nationally representative survey of the U.S. civilian noninstitutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a two-and-a-half year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for two calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when

combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS-HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with oversampling of Hispanics and blacks.

Medical Provider Component

The MEPS-MPC supplements and validates information on medical care events reported in the MEPS-HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the HC. Also included in the MPC are all office-based physicians:

- Providing care for HC respondents receiving Medicaid.

- Associated with a 75 percent sample of households receiving care through an HMO (health maintenance organization) or managed care plan.

- Associated with a 25 percent sample of the remaining households. Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9 or ICD-10 (9th or 10th Revision, International Classification of Diseases) and DSMIV (Fourth Edition, Diagnostic and Statistical Manual of Mental Disorders).

- Physician procedure codes classified by CPT-4 (Current Procedural Terminology, Version 4).

- Inpatient stay codes classified by DRG (diagnosis related group).

- Prescriptions coded by national drug code (NDC), medication names, strength, and quantity dispensed.

- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials.

Insurance Component

The MEPS-IC collects data on health insurance plans obtained through private- and public- sector employers. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, and employer characteristics.

Establishments participating in the MEPS-IC are selected through three sampling frames:

- A list of employers or other insurance providers identified by MEPS-HC respondents who report having private health insurance at the Round 1 interview.

- A Bureau of the Census list frame of private-sector business establishments.

- The Census of Governments from the Bureau of the Census.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and other insurance providers) are linked back to data provided by the MEPS-HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to estimates of employer contributions to group health insurance from the MEPS-IC in the computation of Gross Domestic Product (GDP).

The MEPS-IC is an annual panel survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone follow-up for nonrespondents.

Survey Management

MEPS-HC and MPC data are collected under the authority of the Public Health Service Act. Data are collected under contract with Westat. 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) of the Centers for Disease Control and Prevention provides consultation and technical assistance related to the selection of the MEPS household sample.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports, micro data files, and tables via the MEPS Web site: www.meps.ahrq.gov. Selected data can be analyzed through MEPSnet, an online interactive tool designed to give data users the capability to statistically analyze MEPS data in a menu-driven environment.

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

AHRQ Publications Clearinghouse
Attn: (publication number)
P.O. Box 8547
Silver Spring, MD 20907
800-358-9295
703-437-2078 (callers outside the United States only)
888-586-6340 (toll-free TDD service; hearing impaired only)
To order online, send an email to: ahrqpubs@ahrq.gov.

Be sure to specify the AHRQ number of the document or CD-ROM you are requesting. Selected electronic files are available on the MEPS Web site: http://www.meps.ahrq.gov/.

For more information, visit the MEPS Web site or email MEPSProjectDirector@ahrq.hhs.gov.

Table of Contents

Sample Design of the 2017 Medical Expenditure Panel Survey Insurance Component

Karen E. Davis, MA

Background

The Medical Expenditure Panel Survey Insurance Component (MEPS-IC) is an annual federal survey of employers that is a major source of information on employer-related health insurance in the United States. The survey is sponsored by the Agency for Healthcare Research and Quality (AHRQ) and conducted by the U.S. Census Bureau. It is designed to collect employment-related health insurance information, such as whether insurance is offered and if so, the annual premiums, enrollments, employee contributions, and types of offered plans, deductibles, coverage and copayments. Employer characteristics such as firm size, type of industry, average payroll per employee, and other items are also collected.

The survey was first administered in 1997, with data collected for the entire 1996 calendar year. Each year, a large number of tables of estimates are published on the MEPS website for each annual survey (http://meps.ahrq.gov/mepsweb/data_stats/quick_tables.jsp#insurance). These tables provide estimates at the National, State, and Census geographic division levels as well as for selected metropolitan statistical areas (MSA). Data from the MEPS-IC are only released in aggregate tabular format because of Census confidentiality restrictions. Researchers can apply for permission to use the restricted-access microdata at designated Research Data Centers (RDCs). For more information about these RDCs, see: https://www.census.gov/about/adrm/fsrdc/locations.html.

This report describes the sample design, sampling allocation, and sample selection process for the 2017 MEPS-IC. A glossary of terms related to the MEPS-IC is available at: http://meps.ahrq.gov/mepsweb/survey_comp/ic_ques_glossary.shtml

Sample Design Process Overview

The MEPS-IC is a nationwide sample of private-sector establishments and State and local governments. Data are collected from samples selected from two sampling frames that, together, cover nearly all of the employers in the United States, with the exception of the Federal Government and the U.S. military which are not part of the target population. The two sampling frames are as follows:

Private-sector
The U.S. Census Bureau's Business Register (BR) is a confidential list of private-sector establishments, developed and maintained by the Census Bureau, which is continually updated. It is the source of official Census Bureau figures on the number and employment size of establishments in the United States.

### State and Local Government (Public) Sector
The frame of State and local governments for the MEPS-IC is the Governments Master Address File (GMAF), constructed with units that are eligible from the Census of Governments (COG) and updates from several annual economic surveys. The COG is

conducted every 5 years by the Census Bureau and is updated continually between Census years. For more information about the COG, see:
http://www.census.gov/econ/overview/go0100.html

The two prongs of the survey undergo separate sample selection and estimation processes. The combined sample consists of almost 46,000 employers (see Figure 1).

The overall sampling goal for the MEPS-IC is to produce nationally representative estimates for the private and State and local government sectors separately and combined as well as by State for the private sector and by Census Division for State/local governments. There were several precision goals for the 2017 MEPS survey in terms of relative standard errors (RSE) as shown in Appendix A. Figure 1 below provides an overview of the sampling processes and sample sizes in 2017 while sections on Private-Sector and State and Local Government of this report describe these processes in more detail.

Figure 1. 2017 MEPS-IC Sample Allocation Summary
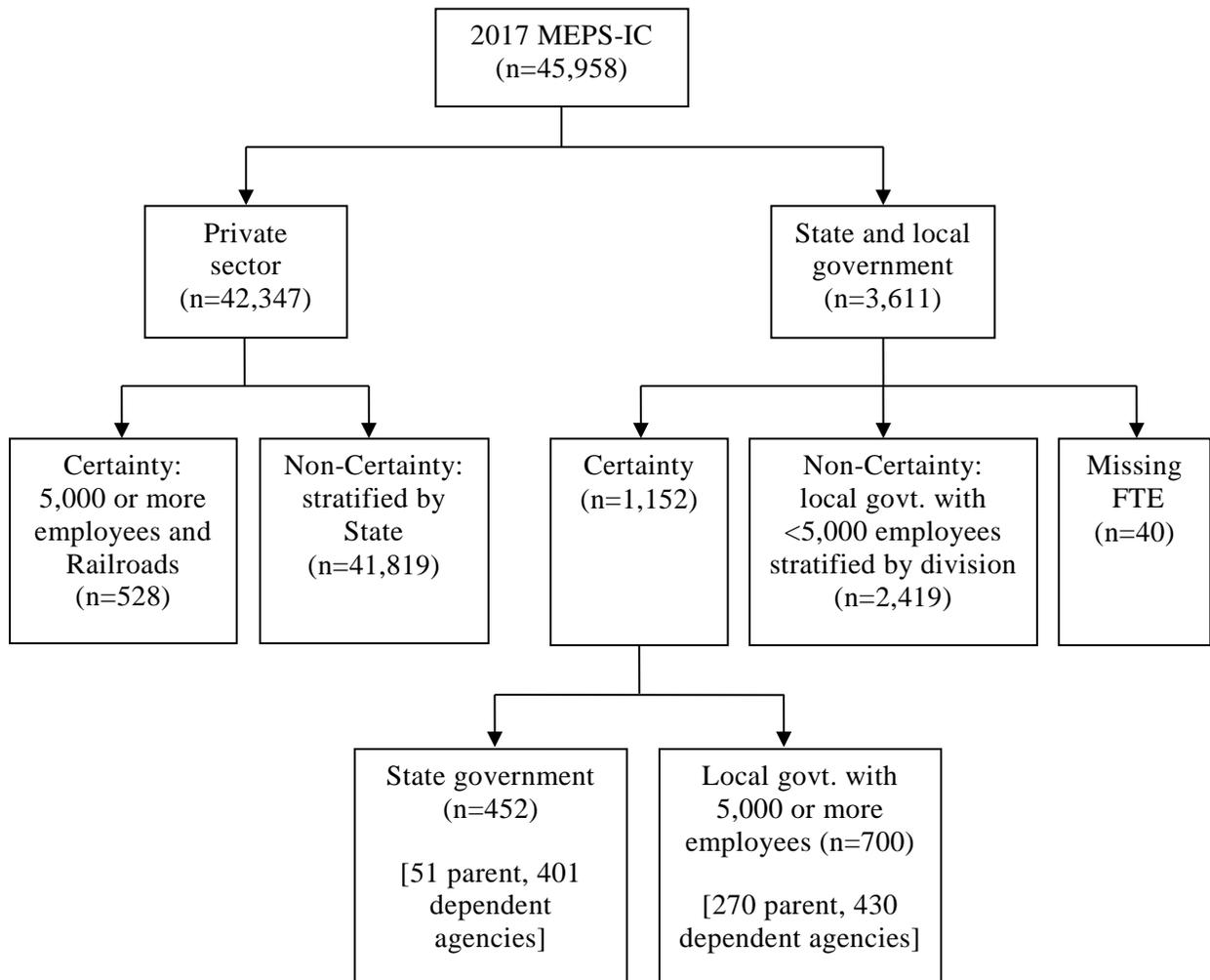
Private Sector

Frame

The private sector frame is created from the Census Bureau's Business Register (BR) and is constructed each year in March, following the timing of payroll imputation processing which is usually not completed until February. For the private sector, an establishment is defined as a particular workplace or location, while a firm is a business entity consisting of one or more business establishments under common ownership or control. In this report, establishments within firms that have more than one establishment are referred to as multi-units while other establishments are referred to as single-units.

For the 2017 MEPS frame, a single-unit establishment was included if its annual payroll was greater than zero in 2016 while multi-unit establishments were included if the annual payroll was greater than zero in 2015. Two different years were used to develop the 2017 MEPS frame because a major change to the frame construction occurred in 2008 when the survey switched from retrospective (with the interview conducted in the calendar year following the survey reference year) to current (with the interview year the same as the survey reference year) (Kearney and Sommers, 2006). This change impacted the choice of data to use to determine whether establishments are in-scope and which data are available to place them in strata. Consequently, the data year used for multi-units is one year older than for single-units because multi-unit imputation processing was not completed at the time of frame construction. There were about 7.3 million private-sector establishments in the U.S. in 2017. Note that for 2017, all large establishments with 5,000 or more employees were selected with certainty.

The following types of establishments on the BR are considered out-of-scope: U.S. Post Offices; private households; public administrations; insurance and employee benefit funds; trusts, estates, and agency accounts; offices of bank holding companies; and offices of other holding companies. Unincorporated self-employed establishments with no employees (SENEs) are excluded from the MEPS-IC frame.

Special processing occurs for railroads and single-unit agriculture production establishments. Railroads are handled in a special way because these data do not correspond to any one State (or site) and are often at the firm level instead of the establishment level. Thus, State-level data for railroads are not available on the Business Register. Because of this, all railroad firms are included in the sample (i.e., treated as certainties). In addition, the negligible number of non-railroad establishments associated with these firms are excluded from the frame. Single-unit agriculture production establishments are temporarily pulled out from the MEPS frame before the private-sector sample is drawn because there are no edits for them on the BR. These establishments are edited separately, known out-of-scopes are removed, and employment is imputed if it is missing or zero using annual payroll data, average quarterly wage factors and other data from the Bureau of Labor Statistics. After the editing process, these agricultural establishments are added back to the MEPS frame in preparation for sampling. On average, about 750 of these cases are sampled each year.

When frame construction is complete, four panels are created where each non-certainty establishment is randomly assigned to 1 of the 4 panels (see section on Private-Sector Sample Allocation and Selection below for definition of "certainty" and "non-certainty" establishments). When combined with the certainty establishments, each of these panels is nationally representative. Multi-unit establishments on the prior year's frame are assigned

to the same panel as the prior year, while single-units and new multi-unit establishments are randomly assigned across the 4 panels. Each year, 2 of the 4 panels are selected for the survey comprised of one new panel and one old panel overlapping the prior year. This strategy helps to reduce the reporting burden for multi-units by reducing their chances of being repeatedly included across years into the MEPS-IC sample.

Sample Allocation and Selection

The private sector sample is drawn at the establishment level, not at the firm level, so it is possible to have more than one establishment sampled from the same firm. There is a certainty stratum which contains establishments with employment of 5,000 or more. All of these establishments in the U.S. are selected and are not part of the State allocation process for the non-certainty sample described below. Railroad establishments are also selected with certainty in their own stratum.

For the non-certainty establishments, the optimal national allocation to States would be to allocate them proportional to the number of establishments within each State. However, for most States this would result in far too small a sample to meet State estimation goals. From experience with past MEPS-IC surveys, it has been determined that a sample of approximately 500 establishments per State yields estimates that meet most State estimation goals using State stratification and allocation processes. To meet State precision goals, an equal size sample could be allocated to each State. An allocation of equal sample to each State would produce State estimates that meet State estimation goals, but would be 50 percent less precise nationally than proportional allocation and would not produce national estimates that meet the precision target. Therefore, a compromise allocation was developed which starts by proportionally allocating about 21,000 sample establishments (based on the assumption of an 80 percent response rate) among the States. The allocation is then augmented for the 42 smallest States so that each of the 11 smallest States receive 495 additional sample establishments and each of the next 31 largest States receive 535 additional sample units. The 9 largest States are not augmented and therefore receive their entire sample allocation from the proportional allocation of the 21,000 units. This allocation results in sampling error for national estimates about 20 percent higher than if the entire sample were proportionally allocated. However, these estimates do meet national and State estimation goals (Appendix A).

Table 1 provides the 2017 MEPS private-sector sample allocation for non-certainties by State. The total allocated sample size is 41,819.

Table 1. Private-Sector Non-Certainty Allocations by State, 2017

State	Allocated Sample Size*	Total Responding
Alabama	787	463
Alaska	672	449
Arizona	726	417
Arkansas	672	399

California	2,748	1,578
Colorado	726	440
Connecticut	725	428
Delaware	672	343
District of Columbia	672	358
Florida	1,122	653
Georgia	792	445
Hawaii	672	365
Idaho	672	432
Illinois	726	419
Indiana	726	482
Iowa	726	498
Kansas	725	455
Kentucky	726	450
Louisiana	726	420
Maine	672	448
Maryland	726	398
Massachusetts	725	440
Michigan	756	475
Minnesota	726	476
Mississippi	672	399
Missouri	725	430
Montana	672	448
Nebraska	672	422
Nevada	672	369

New Hampshire	672	416
New Jersey	1,230	612
New Mexico	672	435
New York	1,437	748
North Carolina	984	643
North Dakota	672	473
Ohio	936	577
Oklahoma	726	454
Oregon	725	485
Pennsylvania	1,147	684
Rhode Island	672	372
South Carolina	726	482
South Dakota	672	485
Tennessee	726	506
Texas	1,837	1,094
Utah	726	514
Vermont	672	503
Virginia	876	526
Washington	755	497
West Virginia	672	459
Wisconsin	726	494
Wyoming	672	462
Total*	41,819	26,694

* Total responding as of April 5, 2018.

After the State sample sizes are determined, the sample is allocated into 14 strata within each State. The 14 strata are defined by a combination of establishment size and firm size.

The 2017 MEPS strata boundaries and allocations are listed in Table 2 below. Note that these stratum boundaries are evaluated periodically and subject to slight modifications in different years.

Table 2. Private-Sector Stratum Boundaries and Non-Certainty Allocations, 2017

Stratum	Firm Size (# of employees)	Establishment Size (# of employees)	Total Allocation Across States
11	1–12	1–3	5,604
12		4–12	6,983
21	13–87	1–25	4,669
22		26–87	4,638
31	88–722	1–18	1,296
32		19–64	1,468
33		65–135	1,428
34		136–272	1,014
35		273–722	704
41	723+	1–20	4,356
42		21–87	3,566
43		88–279	2,946
44		280–924	1,890
45		925–4,999	1,257

A composite of two different allocations based on the Neyman optimal allocation formula (Cochran, 1977) is used to obtain the State-level non-certainty allocation for the ith stratum within each State as follows:

$r_{si} = .11 \, n_{si} + .89 \, m_{si}$

The first allocation is performed as follows based on the standard deviation calculated for the estimated percent of all establishments that offer health insurance:

$$n_{si} = \frac{N_{si}S_{1si}}{\sum\limits_{i=1}^{14} N_{si}S_{1si}} n_s$$

where

N_{si} is the number of establishments in the ith stratum in the sth State,

n_s is the State sample size,

S_{1si} is the standard deviation for the sth State and the ith stratum calculated based on the percentage of all establishments that offer health insurance and,

n_{si} is the allocation to the ith stratum in the sth State.

The second allocation is performed in the same manner but using a different key MEPS-IC estimate (total enrollees) as follows:

$$m_{si} = \frac{N_{si}S_{2si}}{\sum\limits_{i=1}^{14} N_{si}S_{2si}} n_s$$

where

N_{si} is the number of establishments in the ith stratum in the sth State,

n_s is the State sample size,

S_{2si} is the standard deviation for the sth State and the ith stratum calculated based on total enrollees, and

m_{si} is the allocation to the ith stratum in the sth State.

The final allocation, r_{si}, is the weighted allocation obtained by taking the weighted value of the optimal allocations for the two variables. The weighting factors for the final allocation (.11 and .89) were determined based on an evaluation of the best overall balance in precision of estimates for the 2 variables.

Once these allocations are completed, each establishment in a stratification cell is given the same chance of selection equal to

$p_{si} = r_{si}/N_{si}$ where r_{si} is the final allocation within the State.

At this point, in order to reduce the reporting burden on large firms—where a single respondent may sometimes be able to provide the information for more than one establishment owned by that firm, the probabilities are adjusted.

The values of the p_{si}'s for all establishments linked to the same firm on the frame are summed. This yields the number of establishments that are expected to be selected for that firm. For a small number of firms, this expected value is large and potentially a burden for

the responding firms. Moreover, since the insurance offered to employees of establishments within very large firms is often similar, it is more efficient to reduce sample within these firms to both minimize burden and increase sample for other establishments.

To reduce this expected number of establishments, the probabilities of selection are reduced to a level that minimizes response burden using adjustment factors that are based on firm size. To make up for this reduction in sample, the probability of selection for all other establishments in a stratification cell that contains an establishment with a reduced probability of selection is increased (see example in Appendix B). The increase is calculated by the amount necessary to have the sum of the probabilities of selection within the strata equal r_{si}. Once these probabilities of selection are finalized, the allocated samples are selected using systematic sampling. To perform this selection, the file is sorted by State, strata, industry and number of employees. This assures a good balance of establishments within strata.

Prior to 2007, a birth sample was included in the sample allocation, in order to capture any newly created establishments after the frame was constructed but prior to data collection. However, the switch to current year data collection in 2008 eliminated the need for an annual birth sample. While the primary focus for this report is the 2017 survey design, there have also been other significant changes to the sampling design since 2003. A history of the changes to the sample allocations can be found in Appendix C.

The sample sizes for private-sector establishments, reported by single-unit and multi-units, beginning with the 1996 survey can be found at the following link: http://meps.ahrq.gov/mepsweb/survey_comp/ic_sample_size.jsp.

In some years, slight modifications are made to the MEPS-IC to improve various aspects of the survey. For details see Section VIII at the following link: http://meps.ahrq.gov/mepsweb/survey_comp/ic_technical_notes.shtml.

State and Local Government

Frame

The frame of State and local governments for the MEPS-IC is the Governments Master Address File (GMAF), constructed with units that are eligible from the Census of Governments (COG) and updates from several annual economic surveys. The GMAF universe is updated continuously, although a formal and comprehensive update occurs during the COG. The COG identifies and describes all units of governments in the U.S., and provides benchmark figures of public finance and public employment, including how governments are organized, how many people they employ and payroll amounts, and the finances of governments. The COG occurs every five years for years ending in "2" and "7" and the 2017 COG was used for the 2017 MEPS-IC frame. There are also annual surveys, such as the Boundary and Annexation Survey, the Annual Finance Survey and the Annual Survey of Personnel and Payroll (ASPEP), which provide periodic updates to the GMAF. From the survey/census collection period, the data are reviewed and edited as necessary, and the GMAF universe is updated 1.5-2 years following initial collection cycle. A parent government is defined as a State or local governmental entity, while dependent agencies are associated with a parental governmental agency and includes entities such as community colleges, libraries, school boards, etc. The sampling unit for governments is the parent agency along with its dependent agencies (if any). Note that for 2017, all dependent

agencies were sampled for certainty governments (see section on <u>State and Local Government Sample Allocation and Selection</u> below for definition of "certainty" governments). There were about 90,000 State and local governments in the U.S. in 2017. The Federal Government, the U.S. military, and U.S. Post Offices are considered out-of-scope for the survey.

Sample Allocation and Selection

The 2017 MEPS-IC State and local government sample consists of three components: certainties, sampled non-certainties, and sampled missing Full-Time Equivalent (FTE) employment cases. The certainty governments are comprised of the 51 State governments (including Washington, D.C.) and any local government with over 5,000 employees (700 cases in 2017). All certainty cases are assigned a base sample weight equal to 1.0.

The non-certainty government sample covers all other governments (except for missing FTE cases described in the last paragraph of this section below) and is stratified by the 9 Census divisions. The divisions are defined in Table 3 below.

Table 3. Census Division by State

Census Division	States
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York, Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
South Atlantic	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
East South Central	Alabama, Kentucky, Mississippi, Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, Texas
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
Pacific	Alaska, California, Hawaii, Oregon, Washington

A non-certainty sample size of 200 governments is allocated to each Census division for a total of 1,800. To perform the selection using probability proportional to size (PPS) sampling, each government is given a measure of size equal to the square root of its total FTE employment (which includes any dependent agency employment). The selection probability (p_{ij}) for a single government is determined as the total final Census division

non-certainty State government allocation (i.e., 200), times the government's measure of size, divided by the sum of all measures of size for all governments within the Census division on the frame.

$$p_{ij} = \frac{200 * MOS_{ij}}{\sum_{i=1}^{n_j} MOS_{ij}}$$

where

MOS_{ij} is the square root of the non-certainty government FTE employment for the ith government unit in the jth Census division,

n_j is the total number of units in the jth Census division.

The non-certainty government sample within each Census division is selected using a systematic PPS sampling from a file sorted by State, type of government (county, city, township, school district, special district) within the State, and by FTE employment within type of government. For every selected case, a base sample weight equal to the inverse of the selection probability (*p*) is assigned.

Table 4 provides the 2017 non-certainty sample allocations for the public sector.

Table 4. State and Local Government Allocations per Census Division, 2017

Census Division	Selected Sample	Total Sample (parent and dependent agencies)
New England	200	295
Middle Atlantic	200	232
East North Central	200	224
West North Central	200	219
South Atlantic	200	360
East South Central	200	276
West South Central	200	271
Mountain	200	301
Pacific	200	241
Total	1,800	2,419

Finally, it should be noted that cases that have missing FTE employment on the frame are placed into a separate file for processing before the non-certainty sample is drawn. A

systematic sample of 40 cases is drawn from the cases in this file. To perform this selection, the file is first sorted by State, type of government, and total employees within type of government (if available). Every sampled case determined to be in-scope is assigned a base sample weight equal to the number of missing FTE cases divided by 40.

Summary

This report described the sample design, sample allocation, and sample selection processes for both the private-sector and State and local governments within the MEPS-IC. This information is important for researchers using the data who wish to understand its sampling structure. The details presented in this report apply specifically to the 2017 data year. Insurance Component data files are not available for public release; however an extensive series of published tables is available at http://meps.ahrq.gov/mepsweb/survey_comp/Insurance.jsp.

References

Cochran, W.G. *Sampling Techniques*. New York: John Wiley and Sons; 1977.

Davis K. *Sample Design of the 2014 Medical Expenditure Panel Survey Insurance Component*. Methodology Report No. 30. June 2015. Agency for Healthcare Research and Quality, Rockville, MD.
http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr30/mr30.pdf

Kashihara D. *Construction of Weights for the 2011 Medical Expenditure Panel Survey Insurance Component*. Methodology Report No. 28. October 2013. Agency for Healthcare Research and Quality, Rockville, MD.
http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr28/mr28.pdf

Kearney A., Sommers J.P. *Switching from Retrospective to Current Year Data Collection in the Medical Expenditure Panel Survey-Insurance Component*. 2006. In *JSM Proceedings.* Business and Economic Statistics Section. Alexandria, VA: American Statistical Association.

Sommers J.P. *List Sample Design of the 1996 Medical Expenditure Panel Survey Insurance Component*. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 6. AHCPR Pub. No. 99-0037.

Appendices

Appendix A. 2017 MEPS-IC Relative Standard Error Estimation Goals

	Private		State and Local Government	
	National	**State**	**National**	**Division**
Average Premiums	0.005	0.030	0.0075	0.0375
Average Contributions	0.015	0.090	0.020	0.100
Proportions	0.0075	0.300	0.010	0.050

Appendix B. Example of Revised Selection Probabilities for Two Private-Sector Firms

Firm	Selection Probability	Revised Selection Probability
Firm ABC		
Estab #1	0.55	0.34
Estab #2	0.75	0.53
Estab #3	0.75	0.53
Firm DEF		
Estab #1	0.20	0.85
Total	2.25	2.25

Let's say Firm ABC has three establishments. If we sum the selection probabilities in column two for the firm, it yields the expected number of establishments to be selected (2.05) for Firm ABC. However, two establishments may be a response burden for the Firm. Thus we reduce the selection probabilities for all establishments for Firm ABC, and make up for this reduction by an increase for Firm DEF.

Appendix C. History of Changes to the MEPS-IC Sample Allocation

Year	Changes
2003	**Private sector** – The strata within each State were redefined and a separate certainty stratum was created. Logistic regression was used to assign establishments to strata in order to obtain a reduction in variance. http://meps.ahrq.gov/mepsweb/data_files/publications/mr18/mr18.shtml#WithinStates Additional funding due to the dropping of the HC-IC link sample allowed for sufficient sample in every state for the purpose of making state-level estimates. Virginia purchased additional sample for their state to support sub-state estimates. See following link for full list of additional samples purchased by States in earlier years. http://meps.ahrq.gov/mepsweb/survey_comp/ic_technical_notes.shtml#stateestimates **State and local governments** – The nine Census divisions were used as non-certainty strata instead of States.
2004	**Private sector** – Within each State, allocation to the strata was determined separately to avoid assigning to a stratum a sample size that was larger than the number of establishments available within that stratum. Due to budget restrictions, the non-certainty strata sample was reduced across all states by approximately 4 percent.
2005	**Private sector** – The allocation was increased for Alaska and Louisiana for this year only. A total of 770 establishments were added to the sample evenly divided between the two States. The extra sample was allocated across the strata that are less likely to have health insurance or likely to contain only small businesses.
2006	**Private sector** – Budget constraints required an additional reduction of 100 establishments from the total allocation. Also, the one-time increase in the allocation for Alaska and Louisiana was dropped.
2007	Due to the transition from retrospective to current year data collection, there was no survey to collect data for 2007.
2008	**Private sector** – Allocation returned to the original stratification method used prior to 2003, with establishment and firm size classes used for placing establishments into strata. The allocation at the State level was the same as in 2006, and a majority of States had 14 strata. However, smaller States had 8 strata since the strata in these States were collapsed due to small allocations in 1996-2002.

2009-2010	**Private sector –** All States were assigned 14 strata and the strata boundaries were redefined.
2011	**Private sector –** Funding provided for an additional 200 sample cases to be included in the overall sample.
2014	Change in method for calculating standard errors to the Taylor Series method.
2017	**Private sector –** Sampling of all certainty establishments **Public sector –** Increase sample for an additional 700 government units, and sampling of all dependencies for certainty governments.