



Methodology Report #27

Sample Design of the 2011 Medical Expenditure Panel Survey Insurance Component



Abstract

The primary purpose of this report is to describe the survey design, sampling allocation, and sample selection process for the 2011 MEPS Insurance Component (MEPS-IC). This information is important for researchers using the data who wish to understand its sampling structure. Following a brief overview, both the private sector and State and local government sector designs are described. The details presented in this report apply specifically to the 2011 data year, however the appendices include a history of sample allocation changes to the MEPS-IC.

Suggested Citation

Davis, K. *Sample Design of the 2011 Medical Expenditure Panel Survey Insurance Component*. Methodology Report #27. March 2013. Agency for Healthcare Research and Quality, Rockville, MD.
http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr27/mr27.pdf

* * *

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.

Center for Financing, Access, and Cost Trends
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850
<http://www.meps.ahrq.gov>

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, the National Medical Expenditure Survey (NMES) in 1987. Beginning in 1996, MEPS continues 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 (9th 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 two sampling frames:

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

Data from the two 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 health insurance. Since 2000, the Bureau of Economic Analysis has used national 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, 540 Gaither Road, Rockville, MD 20850; 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 through the Internet on the MEPS Web site: <http://www.meps.ahrq.gov/>

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

Contents

SAMPLE DESIGN OF THE 2011 MEDICAL EXPENDITURE PANEL SURVEY INSURANCE COMPONENT	1
Background.....	1
Sample Design Process Overview.....	1
Private sector.....	1
State and local government (public) sector	2
Private Sector.....	3
Frame	3
Private-Sector Sample Allocation and Selection	4
State and Local Government.....	8
Frame	8
State and local government sample allocation and selection	8
Summary	10
References	11
APPENDICES.....	12
Appendix A. 2011 MEPS-IC Relative Standard Error Estimation Goals.....	12
Appendix B. 1996 SENES.....	13
Appendix C. Example of Revised Selection Probabilities for Two Private-Sector Firms.....	14
Appendix D. Private-Sector Industry Codes	15
Appendix E. History of Changes to the MEPS-IC Sample Allocation	16

Sample Design of the 2011 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. Plan 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. Since then, a large number of tables of estimates are published on the MEPS Web site for each survey year (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.

This report describes the survey design, sampling allocation, and sample selection process for the 2011 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 excluded from the sample. 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 that is continually updated. It is the source of official Census Bureau figures on the number and employment size of establishments in the United States. 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. There were about 7 million private-sector establishments in the U.S. in 2011. 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.

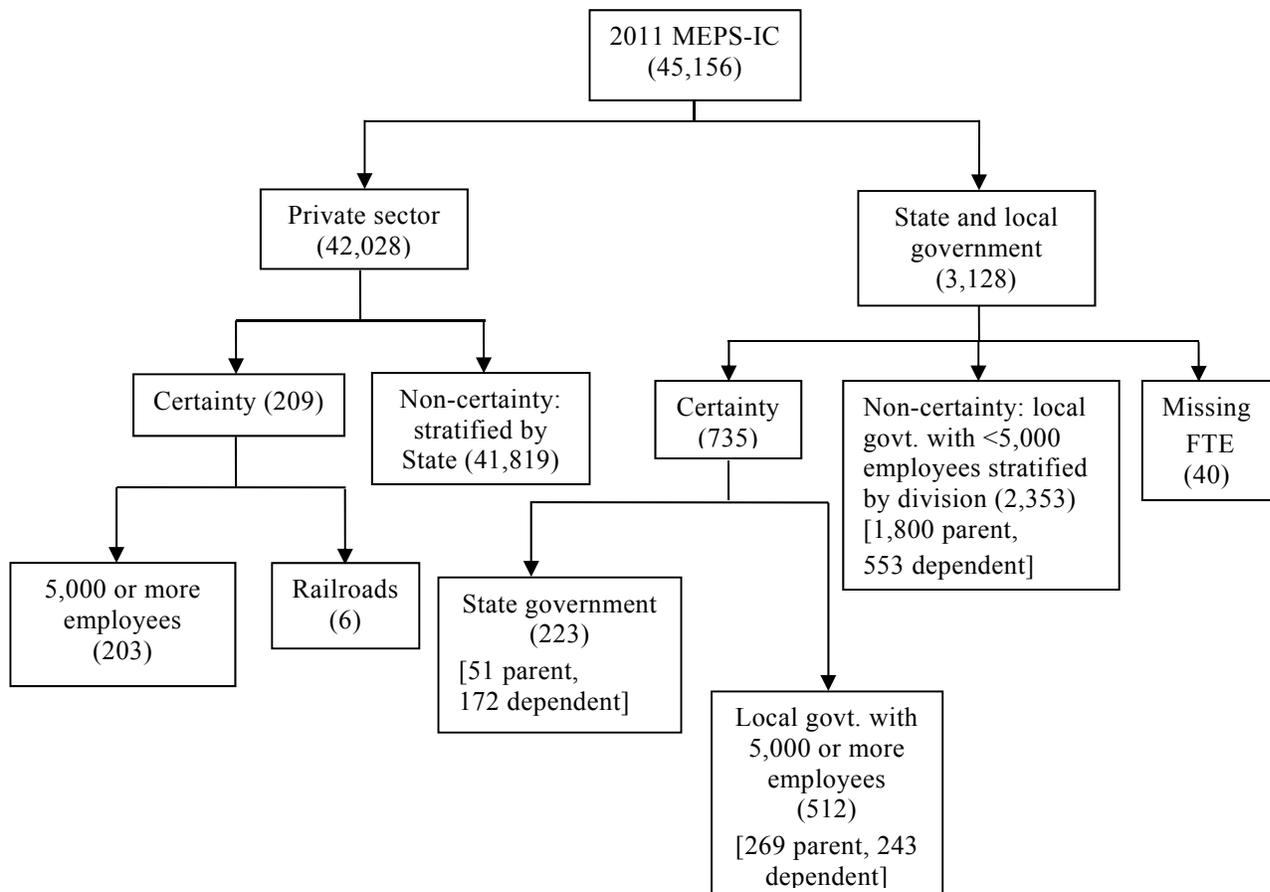
State and local government (public) sector

The frame of State and local governments for the MEPS-IC is derived from the U.S. Census of Governments (COG). The COG is conducted every five years by the Census Bureau and is updated continually between Census years. For the public sector, 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 dependent agencies are not sampled separately. There were about 90,000 State and local governments in the U.S. in 2011. For more information about the COG, see: <http://www.census.gov/econ/overview/go0100.html>.

These two prongs of the survey undergo separate sample selection and estimation processes. The combined sample consists of an independent random sample of about 45,000 employers (see Figure 1). The samples are specifically designed to enable national and State estimates each year.

The overall sampling goal for the MEPS-IC is to produce valid estimates for the private sector for all 50 States and the District of Columbia, State and local governments by Census division, and for the nation as a whole. There were several precision goals for the 2011 MEPS survey in terms of relative standard errors (RSE) as shown in Appendix A. Figure 1 below provides an overview of the sampling process and sample sizes in 2011 while section 2 of this report describes the processes in more detail.

Figure 1. 2011 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 2011 MEPS frame, a single-unit establishment was included if its 2010 annual payroll was greater than zero while multi-units with 2009 annual payroll greater than zero are included. Two different years were used to develop the 2011 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. As a result, subsequent data years are used for single-unit and multi-unit establishments since multi-unit imputation processing has not been completed at the time of frame construction.

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. In 1996 a sample of SENEs was selected using a frame constructed from the Internal Revenue Service. Based on the results for this sample, SENEs were not considered for inclusion in subsequent survey years. Further details on the 1996 SENE sample can be found in Appendix B.

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) and account for about six sampled cases each year. In addition, non-railroad establishments of these firms are excluded from the frame. Single-unit agriculture production establishments are temporarily pulled 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, it is randomly divided into four nationally representative panels and single-unit establishments are randomly assigned to one of the panels. Multi-unit establishments on the prior year's frame are assigned to the same panel as the prior year, while single-units are randomly assigned to one of the four panels. Each year, two of the four panels are selected for the survey with one 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.

Private-Sector 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 and included in their own certainty 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 nine largest States receive their entire sample allocation from the proportional allocation of the 21,000 units. This allocation has an 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 2011 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 per State, 2011

State	Allocated sample size*	Total responding
Alabama	726	565
Alaska	690	575
Arizona	726	552
Arkansas	726	573
California	2,712	1,991
Colorado	726	555
Connecticut	728	555
Delaware	672	483
District of Columbia	674	490
Florida	1,216	891
Georgia	744	545
Hawaii	700	508
Idaho	672	517
Illinois	1,087	804

Indiana	726	554
Iowa	726	569
Kansas	726	569
Kentucky	726	579
Louisiana	744	558
Maine	726	582
Maryland	726	516
Massachusetts	726	556
Michigan	825	632
Minnesota	726	561
Mississippi	726	632
Missouri	726	567
Montana	672	567
Nebraska	726	576
Nevada	744	565
New Hampshire	728	585
New Jersey	772	541
New Mexico	744	584
New York	1,658	1,125
North Carolina	726	577
North Dakota	690	563
Ohio	923	717
Oklahoma	726	558
Oregon	744	572
Pennsylvania	974	751
Rhode Island	674	489
South Carolina	744	577
South Dakota	672	547
Tennessee	726	562
Texas	1,563	1,212
Utah	726	587
Vermont	672	551
Virginia	726	557
Washington	726	565
West Virginia	744	582
Wisconsin	726	570
Wyoming	672	544
Total	41,826	31,998

* Allocated total does not sum to 41,819 due to rounding.

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 2011 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, 2011

Stratum	Firm size (number of employees)	Establishment size (number of employees)	Total allocation across States
11	1-12	1-2	7,273
12		3-5	5,602
13		6-12	5,488
21	13-88	1-18	3,362
22		19-38	3,257
23		39-88	2,099
31	89-708	1-47	2,882
32		48-172	1,630
33		173-708	653
41	709+	1-21	3,365
42		22-90	2,744
43		91-279	1,956
44		280-908	817
45		909-4,999	691

The Neyman optimal allocation formula (Cochran, 1977) was used to obtain the State-level non-certainty allocation for the i^{th} stratum within each State:

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

where

N_{si} is the number of establishments in the i^{th} stratum in the s^{th} State,
 n_s is the State sample size,

S_{1si} is the average standard deviation for the s^{th} State and the i^{th} stratum calculated based on the percent of all establishments that offer health insurance and

n_{si} is the allocation to the i^{th} stratum in the s^{th} State based on establishments that offer health insurance.

After this allocation is completed, a second allocation is performed where a different key MEPS-IC estimate (total enrollees) is used to calculate the average standard deviation.

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

N_{si} is the number of establishments in the i^{th} stratum in the s^{th} State,
 n_s is the State sample size,

S_{2si} is the average standard deviation for the s^{th} State and the i^{th} stratum calculated based on total enrollees and m_{si} is the allocation to the i^{th} stratum in the s^{th} State based on total enrollees.

The final allocation, r_{si} , is the weighted allocation obtained by taking the weighted value of the optimal allocations for the two variables as follows:

$$r_{si} = .44 n_{si} + .56 m_{si}$$

The weighting factors for the final allocation (.44 and .56) were determined based on an evaluation of the best overall balance in precision of estimates for the two variables.

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

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

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 C). 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. MEPS-IC industry codes are listed in Appendix D.

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. In 2011, research was conducted to determine a method for incorporating 200 birth cases into the sample in future surveys. Due to the continued focus by data users on small businesses, as well as health care reform changes, it was recommended that these birth cases be allocated to strata 21, 22, and 23 (see Table 2) and the sample allocated evenly to selected States determined to benefit most from the additional sample.

While the primary focus for this report is the 2011 survey design, there have also been significant changes to the sampling design since 2003. A history of the changes to the sample allocations can be found in Appendix E.

The sample sizes for private-sector establishments, reported by single-units 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 for the MEPS State and local government sample is the Census of Governments (COG) which is conducted every five years and continually updated. The COG is the only source of periodic information that identifies and describes all units of governments in the U.S. It 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 2007 COG was used for the 2011 MEPS-IC frame. The Federal Government, the U.S. military, and U.S. Post Offices are considered out-of-scope for the survey.

State and local government sample allocation and selection

The 2011 MEPS-IC State and local government sample consists of three components: certainties, non-certainties, and sampled missing Full-Time Equivalent (FTE) employment cases. The certainty governments comprised the 51 State governments (including Washington, D.C.) and any local government with more than 5,000 employees (735 cases in 2011). All certainty cases are assigned a 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 nine 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 i^{th} government unit in the j^{th} Census division

n_j is the total number of units in the j^{th} Census division.

The non-certainty government sample within each Census division is selected using a systematic PPS methodology 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 sample weight equal to the inverse of the selection probability (p) is assigned.

Table 4 provides the 2011 non-certainty sample allocations for the public sector.

Table 4. State and local government allocations per Census division, 2011

Census division	Selected sample	Total sample (parent and dependent agencies)	Total responding
New England	200	273	263
Middle Atlantic	199	230	220
East North Central	200	229	228
West North Central	201	222	215
South Atlantic	200	368	334
East South Central	200	283	272
West South Central	199	254	249
Mountain	200	271	233
Pacific	201	223	207
Total	1,800	2,353	2,221

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 sample weight equal to the number of missing FTE cases divided by 40.

Summary

In this report, we described the survey design, sampling 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 2011 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, WG. *Sampling Techniques*. New York: John Wiley and Sons; 1977.

Kearney A., Sommers JP. *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 JP. *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. 2011 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. 1996 SENES

The 1996 MEPS-IC was first administered in 1997 and included an independent random sample of the self-employed with no employees (SENEs). The frame for this sample was the 1994 Internal Revenue Service list of SENEs (Sommers, 1999). The goal was to make a quality national estimate of the SENE population, where quality was defined as a relative standard error of 5.0 percent or less. A national sample of 1,000 SENEs was selected, and divided into five strata based on total income. A one-third loss rate was assumed for nonresponse and out-of-scopes. The target response rate was 75.0 percent. However since many of the sample SENEs were determined not to be legitimate businesses and the response rate was fairly lower (about 40 percent), a decision was made to not sample these cases in future surveys.

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

Firm	Selection probability	Revised selection probability
Firm ABC		
Establishment #1	0.55	0.34
Establishment #2	0.75	0.53
Establishment #3	0.75	0.53
Firm DEF		
Establishment #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 D. Private-Sector Industry Codes

From 1996 to 1999, the industry categories in the MEPS-IC were based on Standard Industrial Classification (SIC) codes. Beginning in 2000, the industries were converted to the North American Industry Classification System (NAICS). Even categories that retained the same name are not comparable for the two coding systems, due to the reclassification of specific businesses from one industry category to another. Making year-to-year comparisons of MEPS data by industries across the 1999–2000 boundary is not recommended.

SIC codes 1996–1999 MEPS	NAICS codes (2000–current)	NAICS sector
Agriculture	Agriculture	11
Fishing	Fishing	11
Forestry	Forestry	11
Mining	Mining	21
Manufacturing	Manufacturing	31,32,33
Construction	Construction	23
Retail trade	Retail trade	44,45
Wholesale trade	Wholesale trade	42
Transportation	Transportation	48,49
Utilities	Utilities	22
Communications	Financial services	52,55
Finance	Real estate	53
Insurance	Professional services	51,54,61,62
Real estate	Other services	56,71,72,81
Services	n/a	n/a

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

Year	Changes
2003	<p>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.</p> <p>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 the 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</p> <p>State and local governments – The nine Census divisions were used as non-certainty strata instead of States.</p>
2004	<p>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.</p> <p>Due to budget restrictions, the non-certainty strata sample was reduced across all States by approximately 4 percent.</p>
2005	<p>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.</p>
2006	<p>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.</p>
2007	<p>Due to the transition from retrospective to current year data collection, there was no survey to collect data for 2007.</p>
2008	<p>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.</p>
2009–2010	<p>Private sector – All States were assigned 14 strata and the strata boundaries were redefined.</p>
2011	<p>Private sector – Funding provided for an additional 200 sample cases to be included in the overall sample.</p>