

TECHNICAL NOTES

Sample Design

These tables were produced using data from the list sample of the Medical Expenditure Panel Survey Insurance Component (MEPS-IC). The MEPS-IC is an annual survey that began in 1997 with data collected for calendar year 1996. The most recently available data is for calendar year 1999. Tables for the 2000 survey will be available in late Summer 2002.

The MEPS-IC is an annual establishment survey that consists of:

- a random sample of business establishments with at least one employee and a sample of government employers (combined they are also known as the List Sample), and
- a set of providers of health insurance to household sample members of the MEPS-Household Component (Cohen, Monheit, Beauregard, Cohen, Lefkowitz, Potter, Sommers, Taylor and Arnett, 1996) (also known as the Household Sample).

In 1996, there was a separate, independent sample of self-employed persons with no employees. This sample was dropped after 1996 due to poor response rates.

The List Sample was selected from two list frames maintained by the Bureau of the Census,

- The Standard Statistical Establishment List (SSEL).
- The Census of Governments.

The SSEL is a list of private sector establishments with at least one employee that is developed and maintained by the Census Bureau. It is derived from administrative records. (Kreps, Slater and Plotkin, 1979) The SSEL is updated on a regular basis as administrative records become available. The MEPS-IC sample for each year is drawn from the SSEL available in the Spring of the following year. This frame contains businesses that existed at the beginning of the sample year and had been supplemented with business births received through the third quarter of that year.

The Governments Division of the Census Bureau produces the Census of Governments once every 5 years and the MEPS-IC survey samples are drawn from the most recently available version. The 1996 MEPS-IC sample was drawn from the 1992 Census of Governments; the 1997, 1998 and 1999 MEPS-IC samples were drawn from the 1997 Census of Governments.

In addition to national estimates, the sample allocation and design of the IC list sample also support reliable State level estimates of:

- ♦ establishment characteristics (for example, the percent of establishments that offer health insurance),

- ♦ employee characteristics (for example, the percent of employees that enroll in health insurance plans) for forty (40) States, and
- ♦ premiums and employee contributions for those enrolled in employer-sponsored health insurance plans.

Survey cost constraints prevent the fielding of a sufficiently large sample to support State estimates for all 50 States and the District of Columbia every year. (For survey purposes, the District of Columbia is treated as a State.) In 1996, estimates were made for the 40 most populous States. Beginning with 1997, the MEPS-IC sample began a rotation of the 20 least populated States so that every State will receive an adequate sample size to make State-level estimates at least once every four years.

The rotation scheme implemented for the twenty States affected for survey years 1996 through 2000 are shown below. "X" indicates the year(s) for which State estimates have been or will be made for that State.

	1996	1997	1998	1999	2000
Arkansas	X	X	X	X	X
Mississippi	X	X		X	X
Utah	X	X	X		X
West Virginia	X		X		X
Nebraska	X		X	X	X
New Mexico	X		X		X
Maine	X	X		X	
Nevada	X	X		X	
Hawaii	X	X		X	
New Hampshire			X		X
Idaho			X		
Rhode Island		X		X	
Montana				X	
South Dakota					X
Delaware			X		
North Dakota					X
Vermont				X	
Alaska		X			
District of Columbia		X			
Wyoming			X		

The State rotation schedule has been modified for 2001 through 2004 to reflect changes in State populations based on the 2000 Census. The rotation scheme planned for the twenty States affected for survey years 2001 through 2004 are shown below. "X" indicates the year(s) for which State estimates will be made for that State.

	2001	2002	2003	2004
Kansas		X	X	X
Arkansas	X		X	X
Utah	X	X		X
Nevada	X	X	X	
New Mexico		X		X
West Virginia		X		X
Nebraska		X		X
Idaho	X		X	
Maine	X		X	
New Hampshire		X		X
Hawaii	X		X	
Rhode Island	X		X	
Montana			X	
Delaware		X		
South Dakota				X
North Dakota				X
Alaska	X			
Vermont			X	
District of Columbia	X			
Wyoming		X		

The MEPS-IC design was developed in steps. The first step was the allocation of sample to the States. An initial sample of 18,500 was allocated to all States proportional to the number of employees in the State. The original proportional allocation number was then increased to a minimum sample size goal of 600 for each of the 40 selected States whose allocation fell below this minimum. The sum of the individual State sample was the desired national total sample of establishments. The lack of minimum sample sizes in States and the District of Columbia is apparent in tables that contain State estimates. In these tables, national estimates and estimates for 40 individual States are given. The remaining States are pooled into an estimate for "States not shown separately".

Next, the State allocations were divided between the private sector and governments. This was done in an iterative manner as follows:

- The Federal and all State governments were determined to be selected with certainty.
- A local government sample was allocated within each State using the local government proportion of the remaining public and private employment within the State. The total of these allocations was used to determine which of the remaining local governments were

large enough to be certainty selections because of their size.

- After removal of this set of local certainty governments, the remaining local governments were allocated a sample based on their proportion of the remaining employment after all government certainties were removed. New certainty governments were determined and the process repeated until no other certainty governments were determined.
- Within each State, the allocation of the sample of the remaining noncertainty governments was determined using the noncertainty governments' proportion of the remaining employment within each State and the remaining sample after the certainty governments were removed.

After the sample was allocated to the public and private sectors, the sample within each sector within each State was allocated to strata. For governments, no further explicit allocation was performed, although the file used for the systematic sample selection of governments was sorted by size within each State.

For the private sector, 14 strata were used within each State. These strata were determined by a cross of the size of the establishment itself and the size of the firm to which the establishment belonged. These two characteristics were used for stratification because:

- Size of firm is correlated with whether an establishment offers health insurance and the characteristics and costs of that insurance.
- Size of establishment is correlated with counts of employees eligible for and enrolled in health insurance.

Allocation to each stratum was determined using variances obtained from the 1994 National Employer Health Insurance Survey conducted by the National Center for Health Statistics (Marker, Bryant and Wallace, 1996), results from previous MEPS-IC surveys, and a Neyman allocation scheme (Cochran, 1977).

The allocations determined the number of cases needed after data collection was complete and non-response accounted for in order to assure adequately small sampling errors. Thus, after allocations were made, the sample sizes were increased to allow for non-response and potential out-of-scope establishments. This assured the final responding sample sizes would match those produced by the allocation process. Selection of the private sector sample within each stratum was accomplished using a systematic selection process. For this selection process the frame was sorted by SIC codes within each stratum (Sommers, 1999). Beginning with the 2001 sample, the MEPS-IC will have completed its industry code conversion from SIC to NAICS and the frame will be sorted by NAICS codes.

Data Collection

Data was collected in two stages. For all sample units, except State and very large local governments, each sample unit was prescreened. The purpose of this step was to determine a

point of contact for data collection and whether or not insurance was offered by this respondent to their employees. If the employer did not offer insurance, a small number of questions were administered and the case was considered a complete respondent. This allowed a quick and inexpensive method to collect the necessary data from the large number of employers who do not offer health insurance to their employees. For establishments that did not offer health insurance to employees, completion of the prescreener finished data collection.

For those establishments that did offer insurance, several brief questions were asked and they were mailed a questionnaire. If they failed to return the mail questionnaire, an attempt to collect the information by telephone follow-up was made. For the purpose of this survey, those who had insurance must have answered key information on their health insurance offering to be considered full respondents. Those that did not provide this information, but were known to offer insurance, were considered partial respondents.

If no contact was made by telephone during the prescreener, a questionnaire was mailed and if not returned, a telephone contact was attempted to collect information. Any employers from this group who responded by mail or telephone were full respondents. Those from this group that were not prescreened, did not return the mail questionnaire and did not respond to follow-up phone calls were classified as non-respondents. For this group, the availability of health insurance for employees at the establishment was unknown.

Estimation

To produce the estimates and their standard errors presented in these tables, weights were created for all responding establishments. Special formulas were used to calculate standard errors. These formulas consider the nature of the sample design. A brief description of these processes is given here.

During the sample design and selection process, each establishment on the frame was given a probability of selection that was dependent on its stratum. These probabilities vary among establishments and assure that the sample sizes in each stratum are equal to that required by the allocation scheme. The inverse of this probability of selection is an establishment's base weight. The use of the base weight and the formula

$$T = \sum_i weight_i X_i$$

provides an unbiased estimate of a total T, if there is no non-response.

Because there is non-response, respondents' weights are adjusted to account for non-response so that these weights, when used with responding establishment data, will reduce the bias attributable to survey non-response. To accomplish this, the sample was divided into cells similar to the original sampling strata and the weights for each respondent in a specific cell were adjusted upward by the same percentage. The sum of the adjusted weights for respondents in these cells was equal to the sum of the base weights for all in-scope sampled establishments in the cell. Because it is assumed that the expected value of all responding establishments in each individual cell defined is equal to that of all the eligible respondents, use of the adjusted weights with respondents should produce the desired unbiased estimates of totals.

After adjustment for non-response, weights were post-stratified (Madow, Olkin, and Rubin, 1983.) using the frame of establishments in business during the last quarter of the year for which estimates would be made to produce control totals. For detailed information concerning construction of weights, see Sommers (MEPS Methodology Report No. 8, November 1999).

Although railroads were included in the sample, the 13 largest railroads were not included in these tables. Employment for these railroads could not be broken down by State so their inclusion would have distorted results for States in which the headquarters of these railroads were located.

Reliability of Estimates

For each table, a corresponding table of standard errors is also provided. Standard errors were produced using the method of random groups. (Skinner, Holt and Smith, 1989.) The method is as follows:

- During the sequential sample selection process, each establishment selected was assigned a number corresponding to its place in the order of selection. These selection numbers were converted to 10 groups numbered 0 to 9 by assigning an establishment to the group determined by the last digit in its selection number. Thus, if the selection number were 73, the establishment would be assigned to group 3. Each group can then be thought of as a subsample similar to the full sample with each unit with a chance of selection into the subsample that was one-tenth its chance of selection into the full sample.
- Using a subsample weights that are 10 times the nonresponse adjusted weights of the full sample, ten subsample estimates, E_i , $i = 1, \dots, 10$ were made in addition to the full sample estimate, E .
- The standard error is calculated as:

$$Std\ Err = \left[\sqrt{\frac{(E_i - E)^2}{90}} \right]$$

Definitions

- Establishment – An economic unit at a single physical location where business is conducted or services or industrial operations are performed.
- Firm – A business with a separate management and legal structure, also referred to as an enterprise. A firm represents the entire organization, including the company headquarters and all divisions, subsidiaries and branches within the company family. A firm may consist of a single establishment or multiple establishments. In the case of a single- establishment firm, the firm and establishment are identical.

- Health insurance plan – An insurance plan that provides hospital and/or physician coverage to an employee or retiree for a single premium.
- Offer health insurance – To make available or contribute to the cost of any health insurance plan for active employees and/or retirees.
- Self-insured plan – A plan offered by employers who directly assume the major cost of health insurance for their employees. Some self-insured plans bear the entire risk. Others insure against large claims by purchasing stop-loss coverage.
- Single coverage – A health insurance plan that covers an individual only. Also known as employee-only coverage.
- Family coverage – A health plan that covers the enrollee and one or more members of their family, as defined by the plan. If more than one family rate was offered, the costs for a family of four were collected.
- Exclusive-provider plan – A plan that requires the enrollee to go to providers associated with the plan for all non-emergency care in order for costs to be covered.
- Any-provider plan – A plan that allows the enrollee to go to the provider of their choice with no cost incentives to use a particular subset of providers.
- Mixed-provider plan – A plan that allows the enrollee to go to any provider but there is a cost incentive to use a particular subset of providers.
- Low-wage employee – From 1996 through 1999, an employee who makes \$6.50 per hour or less. Beginning with 2000, the definition of a low-wage employee is changed to the 25th percentile for all hourly wages in the United States based on data from the Bureau of Labor Statistics. Using this new criterion, the dollar amount will change every year based on the most recent wage data available. For 2000, a low-wage employee is one who makes \$9.50 per hour or less.
- Managed care plan – Either a mixed provider or exclusive provider plan.
- Employee – A person on the actual payroll. Excludes temporary and contract workers but includes the owner or manager if that person works at the firm.
- Full-time employee – A term defined by the respondent. Generally, a full-time employee works 35 to 40 hours per week.
- Part-time employee – An employee not defined as full-time by the respondent.
- Firm size – The total number of employees for the entire firm as reported on the frame.
- Industry – The primary business activity as reported by the respondent. Industry names

were abbreviated in the tables as follows:

agriculture	agric.	utilities	util.
fishing	fish.	finance	fin.
forestry	forest.	insurance	ins.
transportation	transp.	real estate	real est.
communications	commu.		

- Industry group – A set of one or more industries.
- Division (Census division) – The States are grouped in the tables by the following Census divisions:

New England:

Maine
 New Hampshire
 Vermont
 Massachusetts
 Rhode Island
 Connecticut

Middle Atlantic:

New York
 New Jersey
 Pennsylvania

East North Central:

Ohio
 Indiana
 Illinois
 Michigan
 Wisconsin

West North Central:

Minnesota
 Iowa
 Missouri
 North Dakota
 South Dakota
 Nebraska
 Kansas

South Atlantic:

Delaware
 Maryland
 District of Columbia
 Virginia
 West Virginia

South Atlantic (continued):

North Carolina
 South Carolina
 Georgia
 Florida

East South Central:

Kentucky
 Tennessee
 Alabama
 Mississippi

West South Central:

Arkansas
 Louisiana
 Oklahoma
 Texas

Mountain:

Montana
 Idaho
 Wyoming
 Colorado
 New Mexico
 Arizona
 Utah
 Nevada

Pacific:

Washington
 Oregon
 California
 Alaska
 Hawaii

Table Numbering System

- The first level of table access is by the following categories and by year:
 - I. Private-sector data by firm size and selected characteristics
 - II. Private-sector data by firm size and State
 - III. Public-sector data by government type, government size, and census division
 - IV. National totals for enrollees and cost of health insurance coverage for the private and public sectors

- Within each of these categories, tables are subsequently grouped by:
 - A. Establishment-level tables
 - B. Employee-level tables
 - C. Premiums, employee contributions, and enrollment tables for single coverage plans
 - D. Premiums, employee contributions, and enrollment tables for family coverage plans

- Tables within each of these categories are ordered based on their inter-relationships.

This numbering structure also serves as the framework for the MEPSnet/IC interactive tool.

To clarify what each MEPS-IC table is measuring, it will be helpful to use the table (Table 1) provided on the next page. For each of the MEPS-IC tables in categories I, II, and III, this table identifies the denominator for the MEPS-IC table and (where appropriate) the table number previously provided for that table.

Table 1 can also be used to calculate approximate counts for selected tables where percentages are provided. Examples on how to do this are provided in the next section.

Table 1 - Listing of MEPS-IC Table Numbers and Denominators for Tables

Table Number	Denominator for Table		Table Number	Denominator for Table
A.1.				
A.1.a.	A.1.		C.1.	
A.2.	A.1.		C.1.a.	
A.2.a.	A.2.		C.1.b.	
A.2.b.	A.2.		C.1.c.	
A.2.b.(1).	A.2.		C.2.	
A.2.b.(2).	A.2.		C.2.a.	
A.2.b.(3).	A.2.		C.2.b.	
A.2.c.	A.2.		C.2.c.	
A.2.c.(1).	A.2.		C.3.	C.1.
A.2.c.(2).	A.2.		C.3.a.	C.1.a.
A.2.c.(3).	A.2.		C.3.b.	C.1.b.
A.2.d.	A.2.		C.3.c.	C.1.c.
A.2.e.	A.2.		C.4.	B.2.b.
A.2.f.	A.2.		C.4.a.	C.4.
A.2.g.				
B.1.			D.1.	
B.1.a.	B.1.		D.1.a.	
B.2.	B.1.		D.1.b.	
B.2.a.	B.2.		D.1.c.	
B.2.a.(1).	B.2.a.		D.2.	
B.2.b.	B.2.		D.2.a.	
B.2.b.(1).	B.2.b.		D.2.b.	
B.2.c.	B.2.		D.2.c.	
B.3.			D.3.	D.1.
B.3.a.	B.3.		D.3.a.	D.1.a.
B.3.b.	B.3.		D.3.b.	D.1.b.
B.3.b.(1).	B.3.b.		D.3.c.	D.1.c.
B.3.b.(1).(a).	B.3.b.(1).		D.4.	B.2.b.
B.3.b.(2).	B.3.b.		D.4.a.	D.4.
B.4.				
B.4.a.	B.4.			
B.4.b.	B.4.			
B.4.b.(1).	B.4.b.			
B.4.b.(1).(a).	B.4.b.(1).			
B.4.b.(2).	B.4.b.			

Calculation of Approximate Counts

Many of the tables contain percentages of a group of employees or establishments represented by the employees or establishments described on the particular table. For instance, Table I.B.2 gives the percentage of employees who work in establishments that offer health insurance. Table I.B.2.a. gives the percentage of employees who work at establishments that offer health insurance and who are eligible for health insurance. For most tables of percentages, a count of the number of employees or establishments in the cell, with specific characteristics, can be approximated using data, for that cell, from the current table and one or more tables containing the denominator(s) for that cell.

To produce count estimates, one simply multiplies the cell values from the selected table and all of the denominators for that cell. For instance, if one desired an estimate of total establishments that offer health insurance, one can find the percentage of these establishments in Table I.A.2. and determine from the list above that Table I.A.1. contains the value in the denominator of this percentage.

Thus, the estimated total number of establishments that offer health insurance in 1996 is:

$$.529(\text{percentages must be converted to decimals}) \times 5,956,479 = 3,150,977.$$

The first number is from Table I.A.2 and the second from Table I.A.

For some tables, a hierarchical structure exists so multiple tables are used to derive an approximate count. For example, look at Table I.B.2.a. the percentage of employees eligible for health insurance. Table I.B.2. is listed as its denominator and Table I.B.1 is the denominator for Table I.B.2. The values from all three tables, B.1, B.2, and B.2.a must be used to derive an approximate count. Thus, the estimated total number of employees eligible for health insurance is:

$$103,482,267 \times .865 \times .813 = 72,773,387.$$

The numbers are (in order) from Table B.1, Table I.B.2, and Table I.B.2.a.

Basically, one must multiply by a series of denominators until one reaches a table with numbers instead of percents (see the shaded areas of the table on the previous page).

Revision of the Original 1996 Tables

Significant revisions and enhancements were made to the original 1996 tables posted on the AHRQ website. In addition to the significant number of new tables produced and the new table numbering system previously described, additional revisions to the 1996 tables were made based upon:

- ◆ Improved imputation methods using additional data collected during the second year of the survey.
- ◆ Modifications of the weighting methodology to better control for known industry totals.
- ◆ Availability from the SSEL of more precise values for firm size for the time period covered.

AHRQ does not anticipate making revisions of this magnitude in future years. These revisions are critical for anyone wishing to make year-to-year comparisons. The previously issued tables for 1996 should be discarded and not used for this purpose. We are no longer posting them on our website, but they are available upon request.

REFERENCES

Bureau of Labor Statistics. Current Employment Statistics, Most requested series, Series EEU00500005, 1996. Available from URL: <http://www.bls.gov/ceshome.htm>

Cochran WG. Sampling Techniques. New York: John Wiley and Sons; 1977.

Cohen JW, Monheit AC, Beauregard KM, Cohen SB, Lefkowitz DC, Potter DEB, Sommers JP, Taylor AK, Arnett RH. 1996. The Medical Expenditure Panel Survey: a national health information resource. *Inquiry* 33: 373-389.

Kreps J, Slater CM, Plotkin MD. The Standard Statistical Establishment List Program. Washington (DC): United States Bureau of the Census; 1979. Technical Paper No. 44.

Marker D, Bryant E, Wallace L, Yansaneh I. National Employer Health Insurance Survey (NEHIS): draft final methodology report; volume I: statistical methodology. Rockville (MD): Westat, Inc.; 1996.

Madow WG, Olkin I, Rubin DR. Incomplete data in sample surveys, volume 2: theory and bibliographies. New York: Academic Press; 1983.

Skinner CT, Holt D, Smith TMF, Analysis of complex surveys. New York: John Wiley and Sons; 1989.

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.

Sommers JP. Construction of weights for the 1996 Medical Expenditure Panel Survey, Insurance Component list sample. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 8. AHCPR Pub. No. 00-0005.