

SUPPORTING STATEMENT

Part B

Medical Expenditure Panel Survey – Insurance Component 2023-2025

Reauthorization of OMB #0935-0110

Agency of Healthcare Research and Quality (AHRQ)

Version: August 3, 2022

Table of contents

B. Collections of Information Employing Statistical Methods.....	3
1. Respondent Universe and Sampling Methods.....	3
2. Information Collection Procedures.....	8
3. Methods to Maximize Response Rates	9
4. Tests of Procedures	10
5. Statistical Consultants	10

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

The total budgeted sample size for the regular, annual collection is approximately 46,000 sample units before non-response and out-of-scope units. The sample includes all private sector establishments except those noted below, plus State and local governments. The sampling goal is to produce adequate estimates for 1) the private sector for all 50 States and the District of Columbia; 2) State and local governments by Census Division; and 3) the Nation as a whole. This goal is discussed later in this section.

The sample frame is derived from two lists:

- The Census Bureau's Business Register (BR), a list that contains private sector establishments in the United States which employ at least one person. For MEPS-IC purposes, unincorporated businesses that only employ one self-employed person are excluded from the frame (incorporated businesses that only employ one self-employed person are included). The list is derived from tax records, and is continually updated to add newly created establishments (births) and remove those establishments that have closed (deaths). This list contains over 7,800,000 establishments and is very complete, with any under-coverage caused by cases of domestic workers not reported for tax purposes.
- The Governments Integrated Directory (GID), using units that are eligible for the Census of Governments (COG) as the frame. The GID universe is updated continuously, although a formal and comprehensive update occurs during the COG. There are also annual surveys, including the Boundary and Annexation Survey, the Annual Finance Survey and the Annual Survey of Personnel and Payroll (ASPEP), which provide periodic updates. From the survey/census collection period, the data are reviewed and edited as necessary, and the GID universe is updated 1.5-2 years following initial collection cycle. When the 2023 sample is drawn in March 2023, the latest updates to the GID will likely have come from the 2021 ASPEP survey.

Together these two lists cover almost 100 percent of all organizations with at least one employee in the economy, excluding the Federal government.

In order to meet the goals of the survey, the sample is selected with appropriate representation of all important subgroups. These are:

- the sample is allocated to the state and local government and private sectors
- a small sample is set aside for certainty units in both the private sector and state and local governments
- the remaining sample within the private sector is allocated by state and the government sample is allocated by Census Division
- the sample is allocated to strata within each sector.

Allocation to the State and Local Government and Private Sectors

The division of sample between the state and local government and private sectors is based upon past allocations. There are several precision goals for the survey. There are National and State goals for the private sector, and National and Census Division goals for government.

Among the national relative standard error (RSE) goals for the private sector for the survey are the following:

- a .005 RSE for national estimates of premiums

- a .0150 RSE for national estimates of employee contributions
- a .0075 RSE for national estimates of important proportions, such as the percent of employees enrolled in health insurance

Among the national RSE goals for the state and local government sector for the survey are the following:

- a .0075 RSE for national estimates of premiums
- a .020 RSE for national estimates of contributions
- a .010 RSE for national estimates of important proportions

State estimate goals for the private sector are that the RSE for state estimates have errors less than 6 times the similar national private sector goals. Census Division goals for the state and local government sector are that the RSE be less than 5 times the national state and local government goals.

These goals reflect the preference among users for higher quality private sector estimates as opposed to estimates for governments.

Given these goals and the budget limitation for the sample of about 46,000 units, approximately 42,000 sample units are allocated to the private sector and 4,000 to governments. This allocation was based upon the RSEs from past surveys.

Within the private sector, the allocation is broken into 2 parts:

- a small number of approximately 573 large certainty units,
- the remaining sample which is allocated to individual states.

State Allocation for the Private Sector

The optimal national allocation to states would be to allocate proportional to the size of each State. However, for most states this would result in far too small a sample to meet state estimation goals. The 9 largest states under proportional allocation would receive over half the sample and many of the remaining states would have a sample of less than 100 establishments. From experience with past MEPS-IC surveys, it has been determined that a sample of approximately 500 responding 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% less precise nationally than proportional allocation and would not produce adequate national estimates.

The compromise allocation starts by proportionally allocating 17,000 responding sample establishments among the states. The allocation is then supplemented for the 42 smallest states so that each of the 11 smallest states receive about 350 sample establishments and each of the next 31 largest states receive 400 sample units. The 9 largest states receive their entire sample allocation from the proportional allocation of the 17,000 units. This allocation has an error about 20% higher than if the entire sample were proportionally allocated. However, national estimation goals should still be met and the state goals should also be met, with the exception of the 11 smallest states where target quality could be missed by a small degree.

Once the desired responding sample size is determined, the sample allocations are increased to allow for expected nonresponse and out of scope establishments to arrive at the final sample size per state. These sizes are listed in Attachment F.

Private Sector Sample Selection

Once the allocations are complete (excluding the certainty units) for each state, samples are selected for the private sector within each state. The frame to be used for the private sector will be the most recent update of the Business Register (BR). The BR currently contains about 7.8 million establishments. Before final sample allocation and selections, the universe in each state is stratified into 14 stratification cells.

In addition, there is a certainty stratum for each state which contains establishments with projected enrollments of above 5,000 employees. The sample for certainty establishments is allocated for the country as a whole and is not part of the state allocation process.

The table in Attachment G shows the breakdown for all states of the percent of private sector establishments in each stratum and defines the strata boundaries for survey year 2023.

Once these cells are created, the frame within each state is broken into the 14 strata. The breakdown is made by state to allow for the best sample within each state to assure quality state estimates.

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:

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

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_{si} is the standard deviation for the s^{th} State and the i^{th} stratum and

n_{si} is the allocation to the i^{th} stratum in the s^{th} State.

This allocation is calculated based on S_{si} for two different key MEPS-IC estimates (percent of all establishments that offer health insurance and total enrollees) and 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:

$$r_{si} = .01 n_{si} + .99 m_{si}$$

where

m_{si} is the second allocation using total enrollees to calculate S_{si} .

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 create a more efficient sample and 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 burdensome for the responding firms. Moreover, since the insurance offered to employees of different 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 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 employment size. This assures a good balance of establishments within strata.

Allocation and Selection of the State and Local Government Sample

The state and local government sample will use the Governments Integrated Directory (GID) as the frame. This contains approximately 98,000 governments. For this selection there will be only two strata per Census Division. There is a certainty stratum which includes all governments with over 5,000 employees. The certainty stratum for governments accounts for approximately half of all government employment.

The non-certainty governments' stratum contains all other governments. A sample size of 200 governments is allocated to the non-certainty government stratum for each of the 9 Census Divisions.

To perform the selection, each non-certainty government is given a measure of size equal to the square root of its total employment. This increases the sample of smaller governments relative to their total employment. The selection probability for a single government is determined as the total final Census Division non-certainty state government allocation, times the government's measure of size, divided by the sum of all measures of size within the Census Division.

The non-certainty government sample within each Census Division is selected sequentially from a file sorted by state, type of government (county, city, school district, etc.), and employment using a random start.

Sampling of Health Insurance Plans

For private sector establishments, if up to four plans are identified during the prescreener, all are selected for the plan questionnaire. If more than four plans are offered, the three largest plans by enrollment are selected, then a fourth plan is randomly selected. The respondent will receive a preprinted plan questionnaire for each of the selected plans. However, if the case was not contacted during the prescreener and went straight to mail, the survey form asks for the four largest plans by enrollment.

For government establishments, information about every insurance plan offered is collected. Information about all plans offered by governments can be obtained because brochures for the plans are readily available and are provided during data collection. The requisite information about the plans is then abstracted from these brochures. Thus, there is not a disparate burden on governments to provide data for all of their offered plans.

Response Rates and Expected Number of Respondents for the 2023-2025 MEPS-IC

The following table presents the approximate sizes of each of the universes from which sample data are being collected, the approximate expected sample size including out-of-scope, the expected in-scope sample size, the expected number of responses, and the response rates. The response rates are based upon experience from past MEPS-IC data collections.

Projected Response Rates: 2023-2025 MEPS-IC

	Universe	Sample	In-scope	Respondents	Response rate
Private Sector	7,800,000	42,500	40,000	22,500	56.9%
State & Local Government	98,000	3,250	3,150	2,700	86.0%
Total	7,898,000	45,750	43,150	25,200	59.5%

2. Information Collection Procedures

Data Collection

Data collection for the MEPS-IC takes place in three phases: prescreening interview, Eblasts and questionnaire mailout, and nonresponse follow-up. Eblasts are emails sent to sample members with a link to respond to the survey using the internet. The prescreening interview is conducted by telephone. Its goal is to obtain the name and title of an appropriate person in each establishment to whom a MEPS-IC questionnaire will be mailed. Interviewers collect email addresses and also verify mailing addresses and identify businesses that no longer exist, have closed, have merged, etc. Establishments from certain large private firms and state and local governments are not prescreened because they are known to offer health insurance. Due to their size and importance to all Census establishment surveys, the Census Bureau maintains up-to-date contacts with these employers.

For establishments which do not offer health insurance, a brief set of questions about establishment characteristics is administered at the end of the prescreening interview to close out the case. The MEPS-IC establishment and plan questionnaires are mailed to those establishments which, during the prescreening phase:

- were not contacted,
- refused to cooperate,
- were contacted and acknowledged that they did provide health insurance,
- were from large firms or governments specified at the start of collection for mail-only, or
- had no known phone number.

Establishments which do not respond to the initial contacts are mailed a nonresponse follow-up package. Those establishments which fail to respond to the second mailing are contacted for a telephone follow-up using computer-assisted interviewing.

Data for the largest state and local governments and private sector firms, reporting for multiple establishments, are collected using specialized staff and forms. This is done to make the collection process flexible, simple, and as little burden for these important respondents as possible. Sometimes multiple telephone contacts and person visits are used to collect these data. For some collections, abstraction from company records or plan brochures is used if the firm insists on such methods.

Beginning with survey year 2009, a web-based electronic collection tool has been an option for respondents, making it easier to report and also allowing for faster data processing. The Establishment Questionnaire has the following information printed on the front cover:

TO COMPLETE THIS SURVEY ONLINE

Visit: <https://portal.census.gov>

Authentication Code:

Samples of survey materials for the MEPS-IC are included in the Attachments. Included are letters to respondents, the Computer Assisted Telephone Interviewing (CATI) text used for the prescreener, mailed questionnaires, and a definitions list.

Some minor changes will be made in the final materials to reflect current needs. For example, signatures on letters are updated to reflect current management at AHRQ and Census.

Weighting

The private sector and state and local government samples are weighted separately. Beginning with the inverse probability of selection as a base weight, the data are adjusted for non-response. Among other characteristics, the number of plans, size of establishment, size of firm or government unit, industry, and whether the respondent received a mail survey must be considered. Adjustments are made so that the sum of respondents in a cell equals the beginning weighted total for the cells. It would be desirable to adjust for non-response in each state, but cell size may preclude that adjustment in all but the largest states.

Once adjustments for non-response are made, data are poststratified to a set of new control totals. For the private sector, this would be new frame counts from the most recent BR for cells determined by state, industry, size of firm and size of establishment. For governments, poststratification is done to counts provided by the Census of Governments Division. Cells are determined by size of government and state.

Estimation and Accuracy

Estimation is done using sampling weights and variances are calculated using standardized software, such as SAS, using the Taylor series methodology to account for the specialized sampling methods used for the survey.

Certain key variables, such as premiums, contributions and enrollments are imputed when item non-response occurs and values are missing. Selection of donors is accomplished using a nearest-neighbor hot-deck process which chooses the best donor given a set of matching variables and their order of importance. Variables used to match establishments to determine donors are chosen for their correlation with the variable to be imputed, with special care taken to select variables which are also correlated with non-response.

Actual values are derived from donors in various ways. In some cases, ratios are derived from donors and applied to other values the recipient has reported. Great care is taken to maintain consistency of relationships within the data by using ratios and other means that apply data from the donor to data which have been reported by the recipient. The data are also processed in a specific order to assure that variables important to the imputation of variables later in the process are imputed first. For instance, if one must impute the total single premium and total family premium for the same health plan, the values are imputed in order so that the imputed family premium will depend on the single premium that has already been imputed. This is necessary because these values are highly correlated.

Nonresponse Bias Analysis

When an expected unit response rate is below 80 percent, [OMB Standards & Guidelines for Statistical Surveys](#) recommends conducting a nonresponse bias analysis. For the 2023-2025 MEPS-IC we plan to access nonresponse bias in two ways. First, from the sampling frame we will know the size (number of employees), industry category and geographic region of every sampled establishment, allowing us to compare non-responding establishments to responding establishments on these characteristics. If the non-respondents are distributed across these characteristics similarly as the respondents then one might conclude that the non-response may be randomly distributed across the establishments and that the bias due to non-response may be minimized. Second, we will benchmark key survey estimates against estimates from other surveys and administrative data to indicate if any establishment subgroups are in need of additional scrutiny and data collection methods changes.

See Attachment O for details of the most recent nonresponse bias analysis conducted in 2016 using the 2014 data. The results of that analysis indicated potential nonresponse bias among small establishments, establishments in certain industry groups and establishments in certain Census Regions. To address this potential nonresponse bias, each year we monitor these estimates when we benchmark the key MEPS-IC estimates with other surveys.

3. Methods to Maximize Response Rates

To achieve maximum response rates, the following methods are used:

- 1) Perform a screening phone call to identify the best contacts, number of insurance plans and to complete simple cases, such as establishments which offer no insurance.
- 2) Interviewers schedule calls at convenient times for respondents.
- 3) Supervisors regularly attempt to convert refusals.
- 4) Provide respondents the option of completing the survey forms via a secure Internet site maintained by the Census Bureau.
- 5) Mail a self-administered questionnaire upon request.
- 6) Mail a second questionnaire to those not responding in adequate time.
- 7) After an adequate interval, do a telephone follow-up to either remind the respondent or to collect the data.
- 8) For questionnaires returned by mail which fail key edits, perform a callback to verify data.

A further method to improve response, developed to overcome problems encountered in a 1994 predecessor survey and the 1996 MEPS-IC, is a special handling group. This group consists of highly trained analysts, interviewers and statisticians. Their purpose is to collect data from large firms with high burdens. These respondents are given a primary focus early in the data collection process. Each large firm is assigned a collection coordinator. The special handling group can use any reasonable means to accommodate respondents. They can use personal visits and arrange to collect information in formats easiest for the respondents. There are also special collection forms that allow these firms to list plans only once that may be repeated across multiple establishments. Improved methods of collection for large firms allowed the Census Bureau to significantly

improve response rates after the initial 1996 MEPS-IC survey. Another positive effect of these efforts has been a lowering of burden for the larger firms and large governments.

4. Tests of Procedures

This is a renewal of a current survey. As part of general data collection activities, data collection results, interviews, comments made by respondents and estimates are monitored. New data items added to collection are pretested under a separate Census Bureau testing clearance. Because of this monitoring and collection experience, no special pretest is required at this time for the general survey questionnaire.

5. Statistical Consultants

Matthew Thompson
Program Research and Methods Branch
Bureau of the Census
Washington, DC 20233
301-763-4758

Julie Vesely
Chief, Program Research and Methods Branch
Bureau of the Census
Washington, DC 20233
301-763-8099

David Kashihara
Mathematical Statistician
Center for Financing, Access and Cost Trends
Agency for Healthcare Research and Quality
Rockville, MD 20850
301-427-1474

Dr. Joel Cohen
Director
Center for Financing, Access and Cost Trends
Agency for Healthcare Research and Quality
Rockville, MD 20857
301-427-1276

Dr. Sadeq Chowdhury
Mathematical Statistician
Center for Financing, Access and Cost Trends
Agency for Healthcare Research and Quality
Rockville, MD 20850
301-427-1666