

Imputation of Employer Information
for the 1996 Medical Expenditure Panel Survey
Insurance Component

Methodology

Report 10



U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality

Health Care Information and Electronic Ordering Through the AHRQ Web Site



The Agency for Healthcare Research and Quality Web site—<http://www.ahrq.gov/>—makes practical, science-based health care information available in one convenient place.

Buttons correspond to major categories of Web site information, including funding opportunities, research findings, quality assessments, clinical information, consumer health, and data.

The Web site features an Electronic Catalog to the more than 450 information products generated by AHRQ, with information on how to obtain these resources. Many information products have an electronic ordering form and are mailed free of charge from the AHRQ Clearinghouse within 5 working days.

<http://www.ahrq.gov/>

Abstract

The Medical Expenditure Panel Survey (MEPS) is the third in a series of nationally representative surveys of medical care use and expenditures sponsored by the Agency for Healthcare Research and Quality (AHRQ). MEPS comprises four component surveys. The Insurance Component (IC) is a survey of employers, unions, and other providers of health insurance. The IC has two parts. The household sample is linked to sample persons in the MEPS Household Component. It consists of private- and public-sector employers of MEPS respondents, as well as unions and insurance companies that provide insurance to the respondents. The list sample consists of an independent random sample of private-sector business establishments, governments, and the

self-employed with no employees. This report describes the process used to impute values for missing establishment and plan characteristics for the IC in four types of cases: list sample, private sector; list sample, government; household sample, private sector; and household sample, government. The description includes preparation of the data, selection of donors, and the use of donor and other information to create the item for the recipient.

Suggested citation

Sommers JP. Imputation of employer information for the 1996 Medical Expenditure Panel Survey Insurance Component. Rockville (MD): Agency for Healthcare Research and Quality; 2000. MEPS Methodology Report No. 10. AHRQ Pub. No. 00-0039.

Imputation of Employer Information
for the 1996 Medical Expenditure Panel Survey
Insurance Component

Methodology

Report 10



U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality
AHRQ Pub. No. 00-0039
June 2000

The Medical Expenditure Panel Survey (MEPS)

Background

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 also includes a nationally representative survey of nursing homes and their residents. 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 four component surveys: the Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. 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 2½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for 2 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 DSM-IV (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 employers, unions, and other sources of private health insurance. 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 four 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.

- An Internal Revenue Service list of the self-employed.

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 health insurance.

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 followup for nonrespondents.

Nursing Home Component

The 1996 MEPS NHC was a survey of nursing homes and persons residing in or admitted to nursing homes at any time during calendar year 1996. The NHC gathered information on the demographic characteristics, residence history, health and functional status, use of services, use of prescription medications, and health care expenditures of nursing home residents. Nursing home administrators and designated staff also provided information on facility size, ownership, certification status, services provided, revenues and expenses, and other facility characteristics. Data on the income, assets, family relationships, and caregiving services for sampled nursing home residents were obtained from next-of-kin or other knowledgeable persons in the community.

The 1996 MEPS NHC sample was selected using a two-stage stratified probability design. In the first stage, facilities were selected; in the second stage, facility residents were sampled, selecting both persons in residence on January 1, 1996, and those admitted during the period January 1 through December 31.

The sampling frame for facilities was derived from the National Health Provider Inventory, which is updated periodically by NCHS. The MEPS NHC data were collected in person in three rounds of data collection over a 1½-year period using the CAPI system. Community data were collected by telephone using computer-assisted telephone interviewing (CATI) technology. At the end of three rounds of data collection,

the sample consisted of 815 responding facilities, 3,209 residents in the facility on January 1, and 2,690 eligible residents admitted during 1996.

Survey Management

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files.

Printed documents and CD-ROMs are available through the AHRQ Publications Clearinghouse. Write or call:

AHRQ Publications Clearinghouse
Attn: (publication number)
P.O. Box 8547
Silver Spring, MD 20907
800-358-9295
410-381-3150 (callers outside the United States only)
888-586-6340 (toll-free TDD service; hearing impaired only)

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 AHRQ Web site:

<http://www.ahrq.gov/>

On the AHRQ Web site, under Data and Surveys, click the MEPS icon.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality, 2101 East Jefferson Street, Suite 500, Rockville, MD 20852 (301-594-1406).

Table of Contents

Introduction and Background	1
General Process	2
Hot-Deck Process.	3
Process for Each Group of Variables.	4
Government Imputation Process	10
Benchmarks	10
References	11

Appendixes:

A. Establishment Questionnaire.	12
B. Definitions of Selected Variables	18

Imputation of Employer Information for the 1996 Medical Expenditure Panel Survey Insurance Component

by John Paul Sommers, Ph.D., Agency for Healthcare Research and Quality

Introduction and Background

The 1996 Medical Expenditure Panel Survey (MEPS) Insurance Component (IC) is a survey of employers, the self-employed with no employees (SENEs), unions, and insurance companies. The survey is sponsored by the Agency for Healthcare Research and Quality and conducted by the U.S. Bureau of the Census. It is designed to collect employment-related health insurance information, such as premiums and types of plans offered. Information on respondent characteristics, such as size of business, employee characteristics, and industry, is also collected. The 1996 MEPS IC was first administered in 1997; data were collected for the entire 1996 calendar year. Hence, 1996 refers to the data year, not the time of collection.

The sample has two parts:

- The household sample, which is linked to members of the household sample from the MEPS Household Component (HC). It consists of private- and public-sector employers of MEPS respondents, as well as unions and insurance companies that provide them insurance.
- The list sample, which consists of an independent random sample of private-sector business establishments, governments, and SENEs. (An establishment is a single business location, as opposed to a firm, which is a legal entity that can be made up of multiple establishments.)

The IC household sample is defined by the sample design of the MEPS HC (Cohen, 1997) and has persons as sample units. Data are collected from the employers and other insurance providers of the household respondents from the HC. The employers and other providers are proxy respondents for supplemental information on health insurance offered to the household respondent through the employer or other insurance provider. Hence, the probabilities of selection and the corresponding weights for these employers are the same as those of the household sample members and come from the HC design. The data collected in the IC

household sample are combined with other information collected directly from household sample cases, and the two types of data can be analyzed together. For instance, with the use data collected in the HC and the description of coverage in the IC, a person's use of health care can be analyzed relative to the types of health insurance coverage he or she has.

The IC list sample is a random sample of establishments selected only for the IC. Its selection is independent of the HC design (Sommers, 1999). The two IC samples (household and list) are combined for collection purposes. The data collected for both samples are very similar for most cases. To understand the imputation process described in this report, it is essential to understand the type of information collected for each type of case, outlined below.

- List sample, private sector—Information on establishment characteristics, whether insurance is offered, and characteristics of plans offered.
- List sample, government—Information on government characteristics, whether insurance is offered, and characteristics of plans offered.
- Household sample, private sector—Information on establishment characteristics, whether insurance is offered, and characteristics of plans offered, plus a small amount of information concerning the plan selected by the household respondent.
- Household sample, government—Information on government characteristics, whether insurance is offered, and characteristics of plans offered, plus a small amount of information concerning the plan selected by the household respondent.
- SENE—Information concerning the self-employed person's health plan, if any, and some personal demographics.
- Household sample, union—Characteristics of the union, characteristics of plans offered, and which plan the household respondent selected.
- Household sample, insurance company—Plan characteristics of insurance selected by household respondent.

Very similar information is collected for the following cases:

- List sample, private sector.
- List sample, government.
- Household sample, private sector.
- Household sample, government.

A core set of employer information, such as total employment, selected characteristics, and health coverage offered, is collected for all four of these samples. For the two household samples there are a few additional questions about the household respondent.

As with other surveys, individual respondents in the MEPS IC may not answer all the questions presented. As is the custom for most surveys, important items that are missing are completed for all respondents using a process called imputation. This provides the same full set of critical items for each respondent to anyone who wants to perform analyses with the data.

In the case of the MEPS IC, the core of the imputation process is a form of “hot-deck” imputation, by which information for missing items is derived for an individual respondent using information from a similar respondent that has provided the necessary information. The respondent receiving information is called the recipient and the respondent providing information is called the donor. (See Cox and Cohen, 1985, and Kalton and Kasprzyk, 1986.)

This report describes the process used to impute values for missing establishment and plan characteristics for the survey in four types of cases (list sample, private sector; list sample, government; household sample, private sector; and household sample, government) when the respondent returned enough information to be considered usable but did not complete all the questions on the survey form. The description includes preparation of the data, selection of donors, and the use of donor and other information to create the item for the recipient.

The process used for all the private-sector cases for these data items is the same whether the case came from the household or list sample. Likewise, the process used for the data for all government cases is the same. Furthermore, government and private-sector imputation differ only in the selection of the donors to provide

information for missing data. Government cases that require imputation receive information only from other government cases. Private-sector cases receive information only from other private-sector cases.

Because of these similarities, the description of the methodology focuses on the imputation of information for private-sector cases. The question numbers refer to questions on Form MEPS-10, which is shown as Appendix A. After the private-sector process is described, differences in the process for the public-sector data are noted.

The questions listed in this report were considered the critical items for the MEPS IC. They were the only items imputed for survey year 1996.

General Process

The entire process includes more than merely selecting donor values to replace missing values. Many steps are required to create data that are both consistent and logical.

To impute the data for the MEPS IC, the set of items to be imputed was divided into groups. Items within each group are usually related. Groups are composed of either similar types of items or items that require internal consistency and thus must be imputed and edited together. For instance, the first group of items imputed includes numbers of total and part-time employees eligible for insurance and those who are enrolled. Because of size relations among the variables, they were imputed as part of the same process to assure that the values met consistency criteria (e.g., the total number of eligible employees cannot be less than the number enrolled).

The groups of items were also ordered. This is the order of the actual imputation process, chosen to assure overall consistency and correlation within the data records. For instance, if a particular variable was a primary determinant of the value of another variable, it had to be imputed first. Size of premium is an example. The premiums vary by the type of providers that are available to the insured. Can the insured use only a selected group of providers, or can the insured use any provider but some are preferred by the insurance company, or does the plan pay equally for any and all

providers? Thus, before premiums were imputed, type of providers for the plan was imputed. Then the imputed value of the type of providers was used to impute the plan premium, if necessary. This helps retain the relationship between the two variables.

Three steps were usually used to impute items within each group:

- Data preparation—In this step, the data were readied, including such processes as logical edits to fill in results, determination of the set of cases that required imputation for the group of items, and normalization of items to the same basis, as required. For instance, such items as premiums paid might be reported for various time periods—weekly, monthly, or annually. All premium values were converted to a standard time period before imputation.
- Selection of donors—The general process to select donors was the same for all groups. The process is a type of hot-deck imputation developed by Stiller and Dalzell (1997), described below.
- Creation of required values—Once a donor was selected, a value was created for the recipient case. Sometimes this was as simple as using the donor’s item response and placing it in the recipient slot. In other cases, donor information was combined with information from the recipient in more complex ways to create the final value. Values also were edited against other recipient data to create a consistent set of responses.

Hot-Deck Process

A hot-deck method was used to select donor cases to provide information for recipients with items missing. This is a general class of methods by which donors and recipients are divided into prediction classes based on characteristics that predict the value of a given item. For each recipient with a missing data item, a donor in the same class for that variable is selected to provide information for the missing item. For instance, suppose that one needed to produce values of a plan’s total premium when the employer did not respond. If the plan’s premium were predicted by the type of

provider and the State in which the plan was issued, then each time information on a plan’s premium was missing, one could randomly select a plan with the same provider type and State and assign that selected plan’s premium to the plan with a missing value.

The method is based on the assumption, expressed in Kalton and Kasprzyk (1986), that the value of the missing response can be approximated by a regression model based on n parameters with an error ϵ , with $E[\epsilon] = 0$. Thus if y is the missing value, y can be written as:

$$y = \sum_{i=1}^n a_i x_i + \epsilon$$

According to the model, the item being used from the donor has the same expected value as the recipient item being imputed. For the example above, two parameters—State and type of provider—predict y , the respondent’s premium. All plans in the same cell determined by type of provider and State have the same expected premium.

A problem arises when there are no donors in a cell with a recipient. In this case, one needs to determine the best cell from which to select a donor. Many times, in practice, the person performing the imputation simply collapses two or more cells that have similar expected values of the variable to be imputed. The researcher in this case is using predictive mean matching (Little, 1988). This is a method for selecting a value for a recipient from another donor cell with the same or very similar expected value. Of all cells, the donor cell selected has an expected value of the item to be imputed that is the closest to the missing value for the recipient. In terms of expected values, the chosen cell would be termed the “nearest neighbor,” among all cells that contain actual donors, to the cell in which the recipient lies. For the premium example above, if there is no plan with the same provider type and State as the recipient, then the default might be to find a donor in a neighboring State with the same type of providers.

The general method of selecting donors to provide information for imputation in the MEPS IC is similar

to that described in Little (1988). Developed by Stiller and Dalzell (1997), it is a simple automated method to find a very close neighbor to the case with missing data. The method has the following steps:

- Determine the variables to use to define cells for a hot-deck process for a particular item.
- Rank the variables from least important to most important.
- If collapsing must occur, determine the point (going up the list of variables from least important to most important) at which collapsing is no longer acceptable.
- Divide the variables into two groups. Class variables are those that cannot be collapsed. Sort variables can be collapsed.
- Break the file of all respondents that are either donors or recipients for the item to be imputed into separate files determined by the class variables. Recipients within these subsets can receive data only from donors within the same subset. Thus, donors and recipients always have the same values for the class variables.
- Using the sort variables, sort each file defined by the cross of class variables so that the most important variables come first. A recipient's donor generally will be the first donor on the file above it in the sort order. Because the file is sorted from most important to least important, that donor will usually have all sort variables in common. If not, the values that do not match will be the least important variables.

Except for restrictions applied to keep from using the same donor too many times and cases in which there are multiple recipients in a row on the file, the donor selected is the nearest donor on the file to the recipient. (For more detail, see Stiller and Dalzell, 1997.)

Again using the example of missing premium information, suppose it were decided that type of plan was the most important variable and that a plan from a different State would be an acceptable donor if it had the same provider type. Provider type would be a class variable and State would be the sort variable. A separate file of cases with each different provider type would be produced. One would then process through each file individually, continuously updating the values of the

current or nearest donor available in that portion of the file. When a case needing imputation arises, the donor currently available at that point is selected. If there is a large ratio of donors to recipients, as is generally the case, most recipients will be sorted together with donors that match on all or most variables. Because of the sort from most to least important, if the donor and recipient do not match completely, then the matching fails (collapsing occurs) on the least important variables first. The donor currently available will be the best or near-best match in the set.

Process for Each Group of Variables

This section of the report covers each of the groups of variables that were imputed for the MEPS IC. For each group, the variables to be imputed are shown with their questionnaire number. (The questionnaire is given in Appendix A.) Any needed preprocessing is described, sort variables used in the selection of all donors for the individual variables within the group are given, class variables used for imputation are described, and the step-by-step process used to create values for the recipient from the donor information is given.

All logical edits are assumed to have been performed before the imputation takes place. Thus, for instance, if a respondent gave the total number of part-time employees in question D1b as zero and did not fill in how many were eligible or enrolled, these values were automatically set to zero. Because of this assumption, logical edits in process descriptions are not discussed.

Throughout the process, standard definitions of a responding establishment and responding plan are assumed. An establishment was considered a respondent if it answered either that it did or that it did not provide insurance for its employees and, if the establishment did provide insurance for some of its employees, it provided information on at least one of its plans. Responding plans are defined as those for which information was provided on at least one of the following items for the specific plan:

- Type of providers, question B2.
- Gatekeeper required, question B3.

- Purchased or self-insured, question B4.
- Plan enrollment, question B10b.
- Premiums and employer and employee contributions, question B11.

First Group of Variables

The questions in the first group are shown below, along with the applicable questionnaire number (Appendix A).

- D1a Total employees, total eligible, total enrolled.
 D1b Total part-time employees, total part-time eligible, total part-time enrolled.
 D1d Are retirees offered insurance?
 D1d If retirees are offered insurance, which groups are eligible?

Sort Variables

The sort variables are shown in Appendix B. In the order in which they are used, they are: age of firm, industry division group, industry division, Firm Size Class I, and establishment size.

Class Variables

Since this imputation applies only to establishments that offer health insurance, the donor set is immediately limited to the subset of respondents that offer health insurance. Another variable, Firm Size Class 2, is also used as a class variable to subset the population for imputation of this group of variables. This variable divides the firms that own the establishments into two groups based on the total employment of the firm (Appendix B). Using a firm size variable as a sort-order variable and a second firm size variable with fewer categories as a class variable allows the selection of a similar firm size class as a “nearest neighbor.” However, the selection can only be made in the range of firm size classes defined by the variable Firm Size Class 2 (see Appendix B). Using two related variables in this way permits some use of other values of a given variable for the donor but sets a limit on those values by using a class

variable with similar characteristics. This is done often in the MEPS IC imputation process. Industry is also often used as a sort variable and certain groups of industries as a class variable. This means that a donor can have a different industry from the recipient but must belong to the same larger industry group.

Process

If necessary, the total number of employees was taken from the Standard Statistical Establishment List produced by the Census Bureau. This file, which contains a list of almost all private-sector business establishments within the United States, is used as the sampling frame for most Census establishment surveys. It is compiled each year and gives the total employment for the establishment as of the second week in March.

Once the total employment is available, the remaining variables in question D1a are imputed in order, using a separate imputation run for each. To impute total eligible employees, a donor for each recipient is selected from the set of all donors for this group (which is the set of respondents that completed all questions in the group). Two ratios are calculated from the donor: eligible employees to total employees and eligible employees to enrolled employees. If the recipient has reported the total enrolled, the eligible-to-enrolled ratio from the donor is applied to obtain the total eligible. In most cases, when the recipient has not reported the total enrolled, the eligible-to-total ratio is applied to the recipient value of total employment to obtain an estimate of the total eligible. The value of the total enrolled is then imputed by selecting a donor, calculating the ratio of the donor’s total enrolled to total eligible, and applying this ratio to the recipient’s total enrolled eligible employees.

Imputation is continued with the information on part-time employees, D1b. This is calculated by selecting another donor, calculating the ratio of the donor’s part-time to total employees, and applying this ratio to the recipient’s value of total employees. Total eligible and enrolled part-time employees are calculated using a process parallel to that used to calculate total

eligible and enrolled employees: that is, ratios from the donors are applied sequentially, starting with the value of total part-time employees.

Because of the need for consistency in the data, there are minor differences in the process of imputing eligible and enrolled part-time employees and the process of imputing the total number of eligible and enrolled employees. Given that total values have been produced, these values create limits on the sets of values that part-time results can be.

- The two values, eligible and enrolled part-time employees, are both automatically set to zero if the number of part-time employees is imputed as zero.
- The donor set for eligible part-time employees is further limited to only those donors that reported having part-time employees.
- The donor set for enrolled part-time employees is limited to those donors that have eligible part-time employees.
- Each value for part-time employees is limited. These values cannot be set to a value greater than the similar value for all employees (question D1a). For instance, the value of enrolled part-time employees for a recipient is bounded by the value of the total number enrolled, which is imputed earlier in the process. Thus, to impute the number of enrolled part-time employees, one selects a donor and multiplies the recipient's number of part-time eligible employees times the donor's ratio of part-time enrolled employees to eligible part-time employees. This result is then checked against the recipient's total number of enrolled employees. If the imputed number of part-time employees enrolled is greater than the total, then the number of part-time enrolled is decreased to the same value as the total number enrolled.

The final imputation for this group of questions relates to the provision of health insurance to retirees, D1d. This is a two-part question that requires confirmation that the establishment offers health insurance to retirees before the second part of the question is completed. Both parts of the question are imputed in a similar manner. First, from the set of donors using the standard sort and class variables, a

donor is selected to provide a yes/no response to all recipients that require an answer to whether insurance is offered to retirees. Second, for all recipients that answered yes to the first part of the question, including those that had a value of yes imputed for the first part of the question, an answer is imputed to tell what groups of retirees are offered health insurance, under 65 years and/or 65 years and over. Donors for this last step are limited to donors for the group that offers health insurance to retirees.

Second Group of Variables

The question in the second group deals with type of providers (questionnaire number B2 in Appendix A).

Sort Variables

The sort variables, shown in Appendix B, are Census Division, Firm Size Class 1, industry division group, industry division, and State.

Class Variables

The class variables are whether there is a deductible (part of question B14) and Firm Size Class 2.

Process

The process is simple and direct. The donor set is all plans with a response to question B2 (type of provider). Recipients are all responding plans with a missing value of B2. The value for the recipient is determined by direct substitution of the donor value.

Second Group of Variables for Nonresponding Plans

Sort Variables. The sort variables are Census Division, Firm Size Class 2, Firm Size Class 1, industry division group, industry division, and State.

Class Variable. The class variable is the number of plans for which information is requested.

Process. For the standard second group imputation, the value of type of providers, question B2, was imputed

for responding plans. However, type of providers, B2, also can be imputed for nonresponding plans. This imputation is at the establishment level rather than the plan level. When establishments reported for some but not all the plans for which they were required to provide information, this imputation provides an answer to question B2 for the missing plans.

The donor set is the set of all establishments that reported for all the plans required by the survey. An establishment from this group was selected for each recipient establishment that was missing at least one plan. (Note that many establishments reported on some, but not all, plans. Thus an establishment could have reported B2 values for some plans and imputed values for other plans.) A recipient and donor both had to have reported the same number of plans (edited results of question A1).

The plans for both donor and recipient were sorted in this order: question B2 equal to 2, then B2 equal to 3, then B2 equal to 1, then any missing plans. This sets a one-to-one correspondence between the required donor and recipient plans. The plans from the recipient establishment without a value of question B2 were assigned the value of the corresponding donor plan.

Third Group of Variables

The questions in the third group are shown below, along with the applicable questionnaire number (Appendix A).

- B3 Does plan require referral to a specialist?
- B9 Beginning of plan year.

Sort Variables

The sort variables, shown in Appendix B, are Census Division, State, and Firm Size Class 2.

Class Variable

The class variable is type of providers (Question B2).

Process

Donors are all plans collected with the two questions (B3 and B9) completed. Recipients are all responding plans missing at least one of the two questions in the group. There is a separate donor

selection for each question. Thus, a recipient missing both answers would likely have results from two different donors.

Note that, for this group, both donors and recipients are plans rather than establishments. This is a result of the questionnaire design. The questions in Section B of the questionnaire must be answered for every plan offered by an establishment, so answers and the imputation of these results are normally at the plan level. Questions in other sections are answered once for each establishment and thus are worked with at an establishment level.

Fourth Group of Variables

The questions in the fourth group are shown below, along with the applicable questionnaire number (Appendix A).

- B11a Total single premium only.
- B11b Does employer offer family coverage?
- B11b Total family premium only.

Sort Variables

The sort variables, shown in Appendix B, are Census Division, State, Firm Size Class 1, and industry division group. In addition, industry division is used for imputation of total single premium and whether employer offers family coverage, and total single premium is used for total family premium imputation only.

Class Variables

The class variables are provider type and Firm Size Class 2.

Process

The imputation is done in three parts: first, total single premium; second, whether the employer offers family coverage; and last, total family premium. The donor set for total single premium is all plans that have a reported value for both total single and total family premiums (questions B11a and B11b in Appendix A). When a donor is selected for a recipient missing the total single premium, the method of calculating the recipient's premium depends on whether the recipient has a reported value of the total family premium. If

there is a value, then the recipient's single premium is the ratio of the total single to total family premium for the donor multiplied by the recipient's value of the total family premium. If the recipient has no family premium, then the donor single premium value is imputed for the recipient.

The donor set for whether an establishment offers family coverage for the plan is the set of all plans that either stated they do not have family coverage or reported both a total single and a total family premium. The recipient set is those plans for which it is unknown whether family coverage is offered (all plans that did not enter information for any item in question B11b). If a plan reported a premium or contribution, it is assumed that family coverage is offered. If a plan reported that it does not offer family coverage, then it is known that family coverage is not offered. The imputation is a simple yes/no from the donor.

The donor set for the total family premium is the set of all plans that have a reported value for total single premium and total family premium. The recipients are all plans for which family coverage is offered that are missing the total family premium. For this imputation, the file is sorted only by the value of total single premium. The single premium is available for all plans at this time, since this variable has just been imputed. Single premium is the sort variable for family premium because there is a very high negative correlation between the single premium and the ratio of family premium to single premium: the higher the single premium, the lower the ratio of family to single premium. Because of this correlation, the recipient's value of total family premium is calculated as the donor ratio of total family premium to total single premium multiplied by the recipient's value of total single premium. Using the ratio calculation and this particular sort variable maintains the relationship among the values of the single premium, family premium, and their ratio and produces a pair of imputed family and single premiums that parallel the reported pairs of values.

Once values are imputed, a last editing step is done to provide final values for cases that have reported values for either the employee or employer contribution in questions B11a or B11b but not both. The process is the same if either of the contributions is given. If the employer contribution is reported, the employee contribution is the difference between the total value imputed and the employer contribution. If the employer

contribution is greater than the imputed value, the employee contribution is set to zero and the total premium is set to the employer contribution.

Fifth Group of Variables

The question in the fifth group is shown below, along with the applicable questionnaire number (Appendix A).

B4 Is the plan self-insured?

Sort Variables

The sort variables, shown in Appendix B, are industry division group, industry division, Firm Size Class 1, State, same firm.

Class Variables

The class variables are Firm Size Class 2 and "not all HMOs."

Process

To impute this variable requires the imputation of two variables, one at the establishment level and one at the plan level. A variable is created for each establishment that indicates if there are any self-insured plans at the establishment level. The class variable used for this imputation is whether the establishment offers a plan that is not an exclusive provider type plan. This is done because HMOs have been found to be much less likely to be self-insured. The donor set for this variable is the set of all establishments that provided an answer to question B4 (indemnification type) for all their plans. The recipients are those establishments for which the self-insured status cannot be determined. These are respondents that did not report in question B4 that at least one of their plans is self-insured and did not answer question B4 for all their plans.

Once it is established whether an establishment offers at least one self-insured plan, plans for establishments are logically edited. If establishments have a self-insured plan and only one plan, the plan is automatically self-insured. If establishments have no self-insured plans, all plans are assumed to be purchased plans.

For the remaining plans, information on self-insured status is imputed from establishments that have at least one self-insured plan. The donor set is the set of plans from establishments that have at least one self-insured plan. In order to perform this imputation, another class variable is added to the process. This is the provider type for the plan, the answer to question B3. Plans with different answers to question B3 have significant differences in the probability of being self-insured.

Sixth Group of Variables

The questions in the sixth group are shown below, along with the applicable questionnaire number (Appendix A).

- B11a Employer and employee contributions for single coverage.
- B11b Employer and employee contributions for family coverage.

Sort Variables

The sort variables, shown in Appendix B, are industry division group, industry division, Firm Size Class 1, Census Division, and State.

Class Variables

The class variables are Firm Size Class 2 and provider type.

Process

Employer and employee contributions were imputed at the firm level and industry division level. This means that employee contributions in any establishment within the same firm level within the same type of business were assumed to be highly correlated. To accomplish the imputation, first the average employee premium contributions for family and single coverage across all reporting establishments within the same industry division within the same firm level were calculated. For example, if Apex Corporation had both a manufacturing and retailing subsidiary, they were treated separately. If an establishment within the same firm and industry division failed to report this information and other establishments within the same firm and industry division reported, the average employee contribution percentage for the former was assumed to be the same

as for the latter. Thus, for Apex, if a manufacturing establishment was missing employee contributions, this information was imputed, if possible, using the average of other Apex manufacturing establishments.

For firm/industry division combinations that did report for some establishments and could not be completed in this fashion, a donor firm/industry division combination was selected using the standard hot-deck procedure, using the sort and class variables described earlier. In this process, the State for a multi-State firm was set equal to missing.

The donor's average employee single and family contribution percentages were applied to the single and family total premiums for all recipient plans to produce the employee contributions for the recipient plans. Once these were calculated, employer contributions were obtained by subtraction from the appropriate total premium.

Seventh Group of Variables

The question in the seventh group is shown below, along with the applicable questionnaire number (Appendix A).

- B10b Total active employees enrolled in an individual plan.

Sort Variables

The sort variables, shown in Appendix B, are industry division group, industry division, Firm Size Class 1, and State.

Class Variables

The class variables are plan count (number of plans offered to employees) and Firm Size Class 2.

Process

Imputation for this question was limited to establishments that had two or more plans and failed to report for at least two of the plans for which information was requested. All other plans with missing data on total active employees enrolled in a plan can be assigned a value equal to the total enrollment for the establishment, question D1a, less the values for question B10b for the remaining reported plans.

Donors for the process were establishments that offered employees at least two health plans and responded to this question for all required plans. To start the process, donors and recipients were assigned a value called plan count. This was the number of plans the respondent indicated it provided to employees. Plan count was always greater than one for this set.

Selection of donors for recipients was done using the standard method with the sort and class variables given. This means that donors and recipients had the same value of plan count.

Each plan within a selected donor establishment was assigned the proportion of the establishment's total active enrollment that the plan represented (question D1a). If a private establishment offered more than four plans, it was asked to report for the three largest plans plus a fourth selected at random from the remaining plans.

The plans within each donor and recipient establishment were sorted by the values for reported active enrollees. Proportions were assigned from donor to recipient by the sort order.

If the recipient failed to report the active enrollment for all plans, then enrollment for each of the recipient's plans was calculated as the imputed proportion of the establishment's active enrollment. If the recipient reported for some but not all plans, the remaining enrollment in the recipient establishment was distributed among the unreported recipient plans using the relative sizes of the proportions assigned from the donor. For example, suppose there were three plans in an establishment with a total active enrollment of 200 and the respondent gave the active enrollment for the first plan as 100. Further, suppose the donor's three plans represented 50, 30, and 20 percent of the donor's total active enrollment. The remaining 100 active enrollment would be distributed in a ratio of 3 to 2, based on the relative relationship between the two smallest donor plans.

This method attempted to create a distribution of enrollees for the recipient establishment's plans that was similar to that of another establishment. This was done by considering all the plans from both donors and recipients, each with an equal number of plans, in a single step.

Government Imputation Process

In the imputation process for sampled governments, the same data items are imputed as for the private-sector

establishments. The process is similar to that for the private sector. The only differences are the sort and class variables used. For government case imputation, the same sort variables are used for all data groups. These are, in sort order, Census Division, State, and government employment size. No class variables are used to describe the government. The only class variables used are the specialized variables for plans that apply to a particular imputation group. For instance, for Group 4 in the private sector, Firm Size Class 2 and provider type are class variables. For Group 4 government case imputation, Firm Size Class 2 is dropped but provider type is kept. This is because size would refer to government and is not used, but provider type is a characteristic of the plan and these variables are used.

Benchmarks

Review of the results of the imputation and its effects on estimates made with these critical data items was very thorough. After the data were imputed, results were edited to determine if the process had produced inconsistencies in the data. Cell averages were examined before and after imputation to determine if key factors had been missed. During the actual process, problems were found and corrected in this way.

After internal review of the results, estimates were made and benchmarked to outside sources. One key part of this process involved checks against results of the 1994 National Employer Health Insurance Survey (NEHIS), conducted by the National Center for Health Statistics (NCHS). NEHIS was a one-time survey whose purpose, sample design, and questionnaire were similar to those of the MEPS IC (National Center for Health Statistics, 1997). There were some differences in total employment and numbers of establishments, which were easily explained by the number of years between the two surveys and slight differences in the survey frames used. However, the relative numbers benchmarked quite well and had few significant statistical differences. For instance, the estimate of the percent of private-sector establishments that offered health insurance was 52.8 percent for the MEPS IC and 51.6 percent for the NEHIS. Likewise, the percent of workers eligible for health insurance was 80.9 percent in the MEPS IC and 81.7 percent for the NEHIS. A more thorough review of the comparison of the two surveys and some individual numbers from other sources can be

found in the review written by Thorpe and Florence (1999).

References

Cohen SB. Sample design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS Methodology Report No. 2. AHCPR Pub. No. 97-0027.

Cox BG, Cohen SB. Methodological issues for health care surveys. New York: Marcel Dekker; 1985.

Kalton G, Kasprzyk D. The treatment of missing survey data. *Survey Methodology* 1986;12(1):1-16.

Little RJA. Missing data adjustments in large surveys. *Journal of Business and Economic Statistics* 1988;6(3):287-96.

National Center for Health Statistics: Employer-sponsored health insurance: State and national estimates. Hyattsville (MD); 1997. DHHS Pub. No. (PHS) 98-1017.

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.

Stiller J, Dalzell D. Hot-deck imputation with SAS arrays and macros for large surveys. *Proceedings of the 10th Annual NESUG Conference*; 1997. North East SAS Users Group. p. 709-14.

Thorpe K, Florence C. An examination of the Medical Expenditure Panel Survey Insurance Component [unpublished report]. Rockville (MD): Agency for Health Care Policy and Research; 1999.

Appendix A. Establishment Questionnaire

OMB No. 0935-0098: Approval Expires 04/30/98

<p>FORM MEPS-10 (7-7-97)</p> <p style="text-align: center;">U.S. DEPARTMENT OF COMMERCE BUREAU OF THE CENSUS ACTING AS COLLECTING AGENT FOR U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES</p> <p style="text-align: center;">MEDICAL EXPENDITURE PANEL SURVEY (INSURANCE COMPONENT) ESTABLISHMENT QUESTIONNAIRE</p>	
<p>RETURN TO Bureau of the Census 1201 East 10th Street Jeffersonville, IN 47132-0001</p>	
<p>If you have any questions concerning this survey, please call</p>	<p><i>Please correct errors in name, address, and ZIP Code. ENTER number and street if not shown.</i></p>

A FEW IMPORTANT INSTRUCTIONS AND DEFINITIONS

1. For this survey, a **health insurance plan** is defined as providing **hospital and/or physician coverage** for a **single premium** to employees and/or retirees. Exclude extra-cash plans (a specified number of dollars per day in the hospital) or dread-disease (e.g., cancer-only) plans.
2. Coverage could have been purchased from an insurance company, provided by a union or trade association, or self-insured by your company.
3. **Single and family** plans offered by the same insurance company and providing the same level of hospital and physician benefits count as **one plan**.
4. **High and low** options of a plan offered by the same insurance company count as **two plans**.
5. An **HMO** and a **conventional** plan offered by the same insurance company count as **two plans**.
6. If your company operates at more than one location, provide information for the **location on the label** unless otherwise directed.
7. Count **owners and officers** as employees in the enrollment questions if they were eligible for coverage along with the other employees at this location.
8. For the deductibles, copayments, and premiums, **report for typical situations and enrollees**. If cost varies by family size, use a **family of four**. If cost varies by age, provide the information for the average age of your workers.
9. **Estimates** are acceptable if you do not have this information readily available.
10. Provide information for the **pay period that included July 1, 1996** for characteristics such as coverage, premiums, and enrollment. Annual totals, such as costs, should be for **calendar year 1996**, if possible, or for the plan year that included July 1, 1996.

Section A - NUMBER OF PLANS

A1. Did you make available or contribute to the cost of any health insurance plans for your employees or retirees on July 1, 1996? *See instructions 1-5 above for a description of health insurance plans.*

001 1 Yes No 2 No - **If No, go to Section D on page 5.**

How many?

Continue with Section B on page 2.

Section B – PLAN CHARACTERISTICS

B1. On July 1, 1996, what was the name of the health insurance plan with the highest enrollment and its carrier?

If you have received Supplemental Sheets (Form MEPS-10(S)) with plan names preprinted in Question B1, answer only for the preprinted plans. Otherwise, provide data for your 4 largest plans. You may make a copy of the Supplemental Sheet, or Section B of this form, if necessary.

FOR CENSUS USE ONLY

100

012 Name of plan

102 Name of insurance carrier

B2. Indicate the type of providers in this plan.

- 103
- 1 **Exclusive providers** – Enrollees must go to providers associated with the plan except in an emergency. There is typically no cost or a small fixed cost for each physician visit. (For example, HMOs, IPAs, EPOs)
- 2 **Any providers** – Enrollees can go to the physicians of their choice on a fee-for-service basis. The plan does not have any associated providers. (For example, conventional plans, indemnity plans)
- 3 **Mixture of preferred and any providers** – Enrollees can go to a set of "preferred" providers associated with the plan, or providers of their choice. If they go to a non-preferred provider, they face higher costs. (For example, PPOs, POSs)

B3. Did this plan **require** that the enrollee see a primary-care physician in order to be referred to a specialist?

- 104 1 Yes 2 No

B4. Indicate the type of indemnification of this plan.

- 105 1 **Purchased** from an insurance underwriter – Coverage is purchased from an insurance company or other underwriter who assumes the risk for enrollees' medical expenses.
- If purchased, go to Question B6.**
- 2 **Self-insured** – Your company pays the claims from its resources and may charge a premium to employees. The plan may be administered by a *third party*. This type may employ supplemental *stop-loss insurance* to limit unanticipated losses.

For self-insured plans only:

B5a. Indicate if you administered the plan or if you employed a third party.

- 106 1 Self-administered
2 Insurance company or other administrator

b. Did you purchase stop-loss coverage?

- 107 1 Yes 2 No

B5c. Enter this establishment's **total annual cost** of coverage for this plan for the plan year that included July 1, 1996. Include: claims paid, administrative costs, and stop-loss coverage (if any). Include employer and employee contributions.

108 \$.00 *If this is the only plan you offered, also enter this amount in Question C3 on page 4.*

d. Enter the **monthly premium equivalents** (or the COBRA amount if premium equivalents were not calculated) for single and family (of four) coverage for a typical full-time employee. Include the costs entered in B5c. *Also enter this information in Question B11a (single) and B11b (family) – Total premium on page 3.*

109 \$.00 Single coverage

110 \$.00 Family coverage

e. Is the amount entered in B5d –

- 111 1 A premium equivalent?
2 A COBRA amount?

If self-insured, go to Question B7.

B6. Was this plan purchased through a pooling arrangement with other employers such as a multi-employer trust (MET) or a multi-employer welfare arrangement (MEWA)?

- 112 1 Yes 2 No

B7. Was this plan operated by a –

- 113 1 Union 2 Trade Association 3 Neither

114 Name of union or trade association

115 Local number, if a union

116 Name of insurance representative

117 Address (Number and street)

118 City

119 State

120 ZIP Code

121 Telephone number

B8. Did any enrollee receive a direct subsidy or contribution towards any part of the premium (e.g., from a union or government)?

- 122 1 Yes 2 No

B9. In what month did the plan year begin?

Enter a numeric response ¹²³ Month
(e.g., Jan = 01, May = 05).

Section B – PLAN CHARACTERISTICS – Continued

B10a. For this plan, enter the total number of enrollees excluding dependents for this establishment on July 1, 1996.

124

b. Enter the total number of active employees enrolled.

125

c. Enter the number of former employees enrolled through COBRA or other State continuation-of-benefits laws.

126

d. Enter the number of retirees enrolled.

127 Total 128 65 and older

e. Enter the total number of enrollees with single coverage.

129

B11a. Enter this plan's total premium, employer contribution, and employee contribution for a typical full-time employee with single coverage.

If self-insured, enter the monthly premium equivalent from Question B5d on page 2.

130 \$.00 Total premium

131 \$.00 Employer contribution

132 \$.00 Employee contribution

Indicate the premium period Year

133 Week 2 weeks Month Year

b. Enter this plan's total premium, employer contribution, and employee contribution for an enrolled family (of four).

Report for the same premium period as in Question B11a.

If self-insured, enter the monthly premium equivalent from Question B5d on page 2.

134 \$.00 Total premium

135 \$.00 Employer contribution

136 \$.00 Employee contribution

137 Family coverage was not offered

B12a. Did the premiums (not contributions) vary by –

Check all that apply.

- 138 Age?
 139 Sex?
 140 Number of persons (within family coverage)?
 141 Wage or salary levels?
 142 Other? – Specify

099

B12b. Did the amount of the employee contribution (not premium) vary for different employee categories (e.g., full-time, part-time, retiree)?

143 Yes No

B13. Did this plan's premium include either of these services?

Check all that apply.

144 Life insurance 145 Disability insurance

B14. Enter the annual deductibles that enrollees paid out of their pockets before the plan began paying for covered services (using the plan's providers). Many HMO-type plans do not have deductibles.

146 \$.00 Total individual annual deductible OR

Separate deductibles for:

147 \$.00 Physician care

148 \$.00 Hospital care

If the deductible is per overnight hospital stay, report under B15a.

149 \$.00 Total family annual deductible (if applicable)

150 Number of persons – Enter if the plan also specified that the family deductible was met when a number of family members fulfilled their individual deductibles.

151 Plan did not have a deductible

B15a. How much did an enrollee pay for an overnight hospital stay (in a participating hospital, if applicable) after any annual deductible was met?

152 \$.00 → 154 Per day Per stay

OR

153 Percent OR

155 Hospital care was not covered

b. How much did an enrollee pay for an office visit (with a participating physician, if applicable) after any annual deductible was met?

156 \$.00

OR

157 Percent OR

218 Physician care was not covered

B16. What was the maximum amount this plan would have paid for an individual –

a. Over the enrollee's lifetime?

159 \$.00

b. In one year?

160 \$.00

158 No maximum

Section B – PLAN CHARACTERISTICS – Continued

B17. What was the maximum annual out-of-pocket amount for –

a. An individual?

161 \$.00

b. A family (of four)?

162 \$.00

163 No maximum

B18. Indicate which of these services were included in the plan.

Check all that apply.

- 164 Routine mammograms
- 165 Adult routine physical exams
- 166 Routine pap smears
- 167 Office visits for prenatal care
- 168 Adult immunizations
- 169 Child immunizations
- 170 Well-baby care, under 1 year
- 171 Well-child care, 1-4 years
- 172 100% well-baby care
- 173 Chiropractic care
- 174 Other non-physician providers
- 175 Outpatient prescriptions
- 176 Routine dental care
- 177 Orthodontic care
- 178 Nursing home care
- 179 Home health care
- 180 Inpatient mental illness
- 181 Outpatient mental illness
- 182 Alcohol/substance abuse treatment

B19. Could this plan have refused to cover persons with certain preexisting conditions?

183 1 Yes 2 No

Did this happen in 1996?

184 1 Yes 2 No

B20. Could this plan have imposed a waiting period for persons with certain preexisting conditions?

185 1 Yes 2 No

B21a. Is this plan offered in 1997?

186 1 Yes – **If Yes, go to Question B21c.**
2 No

b. If it is not still offered, indicate if it has been –

- 187 1 Replaced with a similar plan
- 2 Replaced by a substantially different plan
- 3 Dropped without offering a replacement – **Go to Section C.**

c. For 1997, enter the single and family enrollments and premiums for this plan or the one that took its place.

Report for the same premium period as in Question B11a on page 3.

188 Single enrollment

189 Family enrollment

190 \$.00 Single premium

191 \$.00 Family premium

Please complete one Supplemental Sheet for each additional hospital/physician plan you offered your employees and retirees on July 1, 1996. You may use photocopies of the Supplemental Sheet or Section B of this form, if necessary.

Section C – GENERAL HEALTH COVERAGE CHARACTERISTICS

C1a. Did you offer **optional** coverage (not included in the basic health coverage) for any of these services in 1996 at an additional premium to the employee?

Check all that apply.

- 192 Dental
- 193 Vision
- 194 Prescription drugs
- 195 Long-term care

b. What was the total amount paid for these coverages in 1996? *Include employer and employee contributions.*

196 \$.00

C2a. Did you impose a waiting period before new employees could be covered by health insurance?

197 1 Yes 2 No

b. What was the typical waiting period?

- 198 1 Less than 2 weeks
- 2 2 weeks to less than 1 month
- 3 1-3 months
- 4 More than 3 months

C3. Enter the total annual cost of coverage for the plan year that included July 1, 1996 for **ALL** hospital/physician plans that you offered **at this location**. *Include employer and employee contributions.*

199 \$.00

Section D – EMPLOYMENT CHARACTERISTICS

D1. Enter the number of employees on your payroll at the location printed on the label for each of the categories below. Report for the pay period that included July 1, 1996. **If you offered health insurance**, also enter the number of employees eligible and enrolled for coverage through your organization. *Include officers and owners. Exclude leased, contract or agency workers.*

a. All employees

200 Total	201 Eligible	202 Enrolled
[]	[]	[]

b. Part-time employees

203 Total	204 Eligible	205 Enrolled
[]	[]	[]

c. Temporary (seasonal) employees

206 Total	207 Eligible	208 Enrolled
[]	[]	[]

d. Were retirees eligible to receive health insurance (other than through COBRA or other continuation-of-benefits laws) on July 1, 1996?

219 1 Yes – *Check all that apply* 2 No

209 Retirees under 65 years

210 Retirees 65 years and over

D2. For the pay period that included July 1, 1996 –

a. Enter the number of women employees 038 []

b. Enter the number of employees 50 years old or older 039 []

c. Enter the number of employees who were union members 040 []

d. Enter the number of employees who earned –

(1) Less than \$6.50 per hour 042 []

(2) Between \$6.50 and \$15.00 per hour 043 []

(3) More than \$15.00 per hour 044 []

D3. How many hours per week must an employee work to be considered full time at your establishment?

041 [] Hours

Section E – COMPANY CHARACTERISTICS

E1. Do you offer any of these fringe benefits?
Check all that apply.

050 Paid vacation

051 Paid sick leave

052 Life insurance

053 Disability insurance

054 Retirement/pension plans

055 Medical Savings Accounts (MSAs)

056 Flexible spending accounts

057 Cafeteria plan –

Enter the average annual value per employee → 058 \$ [] .00

E2. Which of these categories **best** describes your type of ownership?
Check only ONE.

062 1 S Corporation

2 Corporation

3 Partnership

4 Sole Proprietorship

5 Government (Federal, state, or local)

6 Joint venture or cooperative

E3. Is this a nonprofit business?

063 1 Yes 2 No

E4. Which of these categories **best** describes your principal business activity?
Check only ONE.

060 1 Retail trade

2 Personal services (e.g., beauty shops, dry cleaners)

3 Business services (e.g., advertising, computer processing)

4 Other services (e.g., legal and health services)

5 Manufacturing

6 Wholesale trade

7 Finance, insurance, or real estate

8 Transportation, communications, electric, gas, or sanitary services

9 Construction

10 Agriculture or forestry

11 Mining

12 Public administration

E5. How many years has your company been in business? *If you operate at multiple locations, enter the number of years in business for the entire enterprise.*

064 [] Years

E6. Enter the total number of employees your business has at all locations.

034 [] Employees

Section F - TO BE COMPLETED IF YOU DID NOT OFFER HEALTH INSURANCE COVERAGE

F1a. Has your business offered any health insurance as a benefit to the employees or retirees of this location since January 1, 1991?

031 1 Yes No - **If No, go to Question F2.**

b. In what year did your business last offer health insurance coverage to the employees of this location?

032 Last year offered

F2. Did you pay the medical or hospital bills of your employees directly, other than for workers' compensation and/or injuries suffered on the job?

049 1 Yes 2 No

F3a. Instead of providing a health plan in 1996, did you provide a voucher or stipend to your employees which could be used to purchase health insurance?

045 1 Yes 2 No - **If No, go to Section G.**

b. Could this voucher or stipend be used for -

046 1 Health insurance/health care only?
2 Other purposes as well?

c. What was the average value per employee of this voucher or stipend?

047 \$.00 PER → 048 1 Week
2 2 weeks
3 Month
4 Year

500 Remarks

Section G - PERSON COMPLETING THIS QUESTIONNAIRE

212 Name (Please print)

213 Title

Signature

214 Date

215 Telephone number
()

220 Extension

216 FAX number
()

217 E-Mail address

Appendix B. Definitions of Selected Variables

Firm Size Class 1

- 1 if firm employment = 0-5
- 2 if firm employment = 6-24
- 3 if firm employment = 25-99
- 4 if firm employment = 100-999
- 5 if firm employment = 1,000 or more

Firm Size Class 2

- 1 if firm employment = 0-249
- 2 if firm employment = 250 or more

Establishment size

- 1 if establishment employment = 0-10
- 2 if establishment employment = 11 or more

Industry division (Standard Industry Code—SIC)

- Agriculture if two-digit SIC = 01-09
- Construction if two-digit SIC = 15-17
- Retail trade if two-digit SIC = 52-59
- Mining if two-digit SIC = 10-14
- Finance, insurance, and real estate if two-digit SIC = 60-67
- Wholesale trade if two-digit SIC = 50-51
- Manufacturing if two-digit SIC = 20-39
- Transportation, communication, and utilities if two-digit SIC = 40-49
- Services if two-digit SIC = 70-89

Industry division group

- 1 if industry division = agriculture, construction, or retail trade
- 2 if industry division = manufacturing, transportation, communication, utilities, or services
- 3 if industry division = mining, finance, insurance, real estate, or wholesale trade

Firm age

- 1 if age = 0-16 years
- 2 if age = 17 years or more

Census Division

- New England if State = ME, NH, VT, MA, CT, RI
- Mid-Atlantic if State = NY, NJ, PA
- East North Central if State = OH, IN, IL, MI, WI
- West North Central if State = MN, IA, MO, ND, SD, NE, KS
- South Atlantic if State = DE, MD, DC, VA, WV, NC, SC, GA, FL
- East South Central if State = KY, TN, AL, MS
- West South Central if State = AR, LA, OK, TX
- Mountain if State = MT, ID, WY, CO, NM, AZ, UT, NV
- Pacific if State = WA, OR, CA, AK, HI

U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality
2101 East Jefferson Street, Suite 501
Rockville, MD 20852

**BULK RATE
POSTAGE & FEES PAID
PHS/AHCPR
Permit No. G-282**

Official Business
Penalty for Private Use \$300



AHRQ Pub. No. 00-0039
June 2000

ISBN 1-58763-004-4