Methodological Challenges and Innovations of the 1994 National Employer Health Insurance Survey (NEHIS)

Karen Allen and Gail S. Poe
National Center for Health Statistics
6525 Belcrest Road, Room 954
Hyattsville, MD 20782

Key Words: Establishment, Sample, Design, Response

Introduction

During the development of President Clinton's Health Security Act, it became apparent that there were no information systems that could answer many of the important questions relating to employer-sponsored health insurance. As a result, a work group was convened in June of 1993 to develop a new employer survey. The work group consisted of staff from the Agency for Health Care Policy and Research (AHCPR), the Health Care Financing Administration (HCFA), and the National Center for Health Statistics (NCHS). Analysts from the Office of the Assistant Secretary for Planning and Evaluation (ASPE, DHHS), the Department of the Treasury, the Small Business Administration and the Substance Abuse and Mental Health Services Administration (SAMHSA) were also consulted. As a result, a new survey, the 1994 National Employer Health Insurance Survey (NEHIS) was developed and co-sponsored by AHCPR, HCFA, and NCHS. Data collection activities were conducted by Westat, Inc. under the direction of NCHS staff.

Objectives

A major goal of the survey was to collect data on state and national private health insurance spending to be used as input for the National Health Accounts (NHA) maintained by the HCFA. The aim of the NHA is to "identify all goods and services that can be characterized as relating to health care in the nation, and determine the amount of money used for the purchase of these goods and services." _1/ Information from the 1994 NEHIS will be included in the estimate of total spending on private health insurance. The NEHIS will also produce baseline data as of 1993 for evaluating the effects of health care reform and to describe the current employment based health insurance system. State level estimates of employer health insurance are particularly relevant for the states that have already initiated or are planning health care reform measures.

The final objective of the 1994 NEHIS is to provide data for modeling aspects of health care reform. Data items that will be available for microsimulation of employer sponsored health insurance from the 1994 NEHIS are: employee participation rates, the percent of employees with individual coverage, the percent of employees with family coverage, enrollment by type of plan, plan premiums, employer and employee share of premiums, the percent of payroll devoted to health benefits, and the cost of claims.

Sampling unit

The target population was all public sector and private business establishments in the United States. For the private sector, the sampling unit was the establishment. According to the Office of Management and Budget, a private sector establishment is "an economic unit, generally at a single physical location, where business is conducted or where services or industrial operations are performed."_2/ Because many firms have multiple locations that cross state boundaries, surveying establishments rather than firms permits more reliable state estimates of establishments and their employees. State estimates permit interstate analysis of health care policies, analyses of geographic differences in health insurance spending, as well as analysis of the probable effect of proposed health care policy initiatives.

Sampling frames

Three sample frames were used to cover all employers in the U.S.: the Dun and Bradstreet Dun's Market Identifiers (DMI) file was used to sample private establishments, the 1992 Census of Governments (COG) file was used to sample local governments (federal and state governments were included in the sample with certainty), and a national sample of self-employed individuals with no employees (SENE) from the 1993 National Health Interview Survey was also included. All three samples were screened to eliminate duplication in the frames. The sample allocations of the DMI and COG samples were designed to support both employer and employee estimates.

Data Items

The key data items collected on this survey included: For each establishment - total employee counts and types of employees, including full-time, part-time, and seasonal

workers; other establishment characteristics, such as employee wages and company payroll; availability of company-sponsored health insurance and eligibility requirements; and numbers of employees, former employees and retirees eligible and enrolled in a health insurance plan offered by the establishment.

Data items obtained about individual health insurance plans included the name and type of plan, whether self-insured or fully insured, single and family coverage, annual deductibles, amounts of co-payments and co-insurance, plan benefits, and employee and employer health insurance premium contributions.

Beside these items, the NEHIS also obtained additional cost information, including administration costs and claims paid and more detailed information about the specific services covered by the plans.

Types of Establishments

Three types of establishments were defined in the NEHIS; single establishment firms or SEFs, multi-establishments firms or MEFs and self-employed businesses with no employees or SENEs. A SEF was an organization or company with just one sampled location. For example, if a company had three locations but only one was selected for the NEHIS it was defined as a SEF. If, on the other hand, a company had three locations and two or all three were sampled, the establishment become a MEF. SENEs were people who were self-employed and had no other employees.

NEHIS Challenges

The NEHIS presented many challenges which included: 1) selecting the most appropriate sampling frames to maximize coverage, 2) sampling for state and national estimates, 3) tracking and updating business movement, 4) selecting the optimum data collection period, 5) selecting the best data collection methodology, 6) encouraging establishments to participate, 7) identifying the best respondent, and 8) identifying and interviewing multi-establishment firms.

1-2. Selecting the most appropriate sampling frame to maximize coverage by using three frames, the DMI, COG, and NHIS and sampling for state and national estimates by using the establishment as a sampling unit is discussed earlier in this paper in the sections "Sampling unit" and Sampling frames."

3. Tracking and updating business movement

Establishment surveys typically use commercial lists of businesses as a sampling frame. Lists "have the usual inaccuracies such as duplication, incompleteness, errors in classification of units and presence of extraneous units." _3/ Since the NEHIS used the DMI list frame for sampling private establishments, it was necessary to trace the sample businesses that could not be located at the address given on the DMI frame to determine if they had moved or were out-of-business. If they had moved, follow-up was needed to obtain the interview. It is common experience to find a substantial proportion of the sample out-of-business or moved-and-not-locatable even after follow-up. _4/ About 20% of the NEHIS sample was determined to be out-of-scope after tracking and updating the sample file.

4. Selecting the best data collection period

An important decision in conducting the NEHIS was to determine the best time to collect the data. As with household surveys, the heaviest summer vacation periods, July and August, and the holidays, such as Thanksgiving and Christmas, were times we wanted to avoid for the NEHIS data collection. Another time we wanted to schedule our data collection around was employers' health insurance open enrollment season. Since the NEHIS is an employer-sponsored health insurance survey we knew that the persons in the establishments who would be most knowledgeable about the health benefits of their employees would be very busy during that time and would not likely be able to spend an hour or two answering questions. We had hoped to begin data collection in March and end in early July; however, due to schedule delays the NEHIS was not fielded until April, 1994 and continued through December, 1994.

5. Selecting the best data collection methodology

Because mail surveys produce unacceptably low response rates and personal interviews would have been too expensive, the decision was made to use Computer Assisted Telephone Interviewing or CATI to collect the NEHIS data. The CATI methodology provided the NEHIS with a number of distinct advantages and features which included:

a. Maximum flexibility for conducting interviews at business establishments.

Given that interviews were being conducting with persons at their place of work during business hours, interviews needed to be scheduled "at the respondent's convenience." This translated into often having to schedule callbacks to conduct interviews on specific days and times. The CATI system developed for this survey featured a callback scheduler whereby calls would be placed at these specified

times and interviews would be conducted by the next "available interviewer."

In addition to being able to schedule callbacks at the employer's convenience, the system needed to be flexible enough to also allow interviewers to stop interviewing with little advance notice. For example, when the respondent would say "Oh, I have to go now, a customer has just come in the store. Can you call me back later?", the system needed to be felixable enough to handle these situations.

A system that would allow interviewers to talk with more than one respondent and keep track of who was interviewed, when they were interviewed, and what questions they answered was also needed. Frequently, and especially in large firms, there were many "best" respondents depending on whether the questions were related to the establishment's characteristics, such as employee counts, or health insurance plan characteristics, such as premiums or services covered.

b. Efficient Management of NEHIS sample

We also needed a system that could manage the large NEHIS sample and assign cases in such a way that would conform to our limited data collection field period. With CATI literally thousands of sample cases could be in the process of being contacted and interviewed at any given time. Initially, the NEHIS required a system which could generate about 51,000 completed interviews in about four months.

c. Data quality

The CATI is also ideal for administering complicated questionnaire instruments, which the NEHIS has. Since the appropriate questions are selected automatically for the interviewers, errors that result from interviewers following incorrect skip patterns and asking inappropriate questions, or missing questions altogether are greatly reduced. Also, certain kinds of data entry or respondent reporting errors can be flagged at the time the error is made, providing a second opportunity for interviewers to re-key the answer or re-ask the question.

d. Timeliness

Having the questionnaire on a CATI system also minimizes coder data entry errors and thus reduces the time between data collection and the date when the data tapes are ready for release.

6. Encouraging cooperation and participation

To promote trust and enhance the credibility of the NEHIS, we solicited endorsements from unions, trade associations and professional organizations. Endorsements representing private industry were obtained from the Business Roundtable, the Society of Professional Benefit Administrators, the National Association of Manufacturers, and the Health Insurance Association of America. To encourage participation of the self-employed segment of the work force, the National Association of Self-Employed Businesses gave its endorsement. Providing endorsements for the public sector and unions were the National Education Association and the American Federation of State, County and Municipal Employees.

An important factor in obtaining cooperation in telephone surveys is the advance letter. The NEHIS advance letter was sent to the sampled establishment a week or two before the actual interview to notify them about the survey. In addition to the listing of endorsing organizations, the advance letter included a description of the NEHIS, the authority under which the data were being collected, and a provision of confidentiality. We also included in the advance letter a description of the types of information we would be asking for and a list of documents and records that might help them answer the questions.

7. Identifying the best and most knowledgeable respondent

An important objective of the prescreening activity was to obtain the name and title of the person at the sampled establishment who knew most about their health insurance benefits. Obtaining this person's name also enabled us to earmark a specific person to whom we could send an advance letter describing the NEHIS. A study by Paxson, Dillman, and Tarnai reports that the response rates for surveys without an individual named averaged 40% , compared to 72% for those survey mailings addressed to a named individual. _5/ By "personalizing" the NEHIS letters we hoped to foster a sense of cooperation and interest among respondents which would ultimately improve response.

Although prescreening was very helpful in identifying at least one knowledgeable person in the establishment, the kinds of information being obtained from employers often is not available from one individual. In many establishments, information about employee counts versus health insurance plan provisions could only be obtained by interviewing different employees.

While respondents might be willing to try to answer all the questions, we wanted to have a data collection system which enabled interviewers to identify the "most knowledgeable" person or persons to call back if necessary, rather than "get the information" from anyone willing to answer the questions. The flexibility of CATI allowed the NEHIS data collection instrument to be designed whereby the interview could be conducted in segments with multiple respondents, at different times, and with different interviewers. In other words, CATI enabled interviewers to interview the most knowledgeable person for different sections of the NEHIS questionnaire. So-called "gates" preceded most sections of the questionnaire whereby interviewers could establish whether the current respondent would be able to answer the next set of questions or another respondent would be preferable. The NEHIS CATI instrument also contained a Questionnaire Management Screen which allowed interviewers to switch respondents at these "gates" as well as to alter the programmed flow of various sections of the questionnaire.

8. Identifying and interviewing multi-establishment firms (MEFs)

One of the biggest challenges we faced with the NEHIS was determining which businesses in our sample had more than one establishment and how to handle them. We learned from other employer surveys how important it was to identify MEFs prior to the interview. What happens frequently is that the health benefits for establishments in a MEF are administered at the corporate or subsidiary level. Also, more often than not, MEFs can only report aggregate information for the firm and not for specific establishments. When an interviewer calls a sampled MEF establishment, frequently he or she is told to call headquarters for the information. When a different interviewer calls another sampled MEF establishment she is also told to call headquarters. Obviously, it becomes very frustrating for the same respondent to be called by several different interviewers for the same information. By prescreening for MEFs, we were able to conduct the interview at the corporate or subsidiary level, for all sampled MEF establishments whose health benefits were administered at that level. Identifying MEFs prior to MEF data collection was accomplished by using linking information from the DMI file and through alpha matching on establishment names.

In addition to avoiding alienating MEF respondents by repeated phone calls for the same information, prescreening MEFs also allowed us to modify our data collection procedures to accommodate the special needs of larger firms. As mentioned earlier, frequently MEFs could only answer the questions about employee health benefits for all the MEF establishments combined. Because the sampling unit for the NEHIS was the establishment, sections were added to the CATI questionnaire to address the problem of reported aggregate data. For most MEFs this worked. However, for the very large MEFs, or "Mega-MEFs", special procedures were developed for interviewers to collect data off-line and have the data keyed in later.

Respondent burden was another issue with the MEFs because, at times, there were many establishments sampled as part of a MEF and many health insurance plans offered. In order to reduce respondent burden we subsampled five health insurance plans in all establishments that offered more than five plans. Although we collected detailed plan information for only five plans, we did obtain a complete enumeration of all the plans and types of plans offered by the firms. In an effort to reduce respondent burden for the MEFs even more, we did subsampling of both plans and establishments in the largest MEFs.

Refusal conversions were especially important for the MEFs because of the many establishments that were part of a MEF. If an interview was refused at the headquarters or subsidiary level, data for all the establishments whose health benefits were administered at that level were lost. Consequently, we placed considerable emphasis on trying to convert these refusals. More often than not, this task was delegated to Westat supervisors or managers. If a supervisor or manager was unable to convert a MEF refusal, we, as a representative of the Federal Government, made calls to the MEFs. This personalized attention frequently paid off, especially with the very large MEFs.

Evaluation projects:

To help us measure the quality of NEHIS data we have several evaluation projects that are currently ongoing.

1. The Record Check Study is designed to abstract data from health care plan brochures obtained from a sample of NEHIS respondents. Specification error can be measured directly by checking survey responses against (administrative) records. _6/ In addition to comparing data abstracted from the health care plan brochures with NEHIS survey responses, this study will also help us to evaluate the cost and feasibility of this method of data collection, as compared to CATI data collection.

2. The purpose of the cognitive reinterviews is to better understand data quality issues in the NEHIS, to help with data interpretation, and to improve future employer-sponsored health insurance surveys. In-depth personal interviews to obtain direct feedback about their NEHIS experiences were conducted with about 20 Washington area NEHIS respondents.

3. Similar to the cognitive reinterviews, a small number of businesses were re-contacted by telephone by Westat project staff. The purpose of this project was to evaluate the accuracy of their original responses, to learn how respondents interpreted the questions, and to determine whether their responses were consistent with NEHIS concept definitions and critical data items determined to be problematic. Again, this information will be extremely useful is designing future employer-sponsored health insurance surveys.

4. The purpose of the Union and Professional Association Data Collection Activity was to contact unions, associations, brokers, insurance carriers or Third Party Administrators (TPAs) for employers who did not maintain health insurance records for employees who were enrolled in union or association health plans. It became apparent during NEHIS interviewing that many employees were covered under these plans and that employers knew very little about them. The purpose of this activity was to determine if this information could be obtained directly from the unions and professional associations.

5. Approximately 9,000 NEHIS sample cases were classified as single-establishment firm "dead end" cases. These were cases where a valid telephone number could not be obtained or there was no indication that the case was still in existence as a business. Westat and government personnel conducted an informal follow-up study of approximately 50 of these cases in Maryland. In almost all of the cases, the business listed on the Dun and Bradstreet file was not found. This gave us some reassurance that the "dead end" cases did not contain a large proportion of businesses that should have been included in the NEHIS.

References

1. Rice D, Cooper B, Gibson R. U.S. National Health Accounts: Historical perspectives, current issues, and future projections. In: Levy, E. Ed. La Sante Fait Ses Comptes (Accounting for Health). Paris. Economica, 1982.

2. Statistical Policy Office, Office of Information and Regulatory Affairs. Quality in Establishment Surveys. Office of Management and Budget. Statistical Policy Working Paper 15(11). July, 1988.

3. Hidiroglou MA, Srinath KP. Problems associated with designing sub-annual business surveys (220). In: American Statistical Association proceedings of the section on survey research methods. Atlanta, Georgia: Annual Meetings of the American Statistical Association. 1991.

4. Hirschberg DA, Nisselson H. An efficient new dual frame sample design for surveys of small business (662). In: Proceeding of the international conference on establishment surveys. Buffalo, NY: International Conference on Establishment Surveys. 1993.

5. Paxson C, Dillman D, Tarnai, J. Methods for improving response to establishment surveys (496). In: Proceeding of the international conference on establishment surveys. Buffalo, NY: International Conference on Establishment Surveys. 1993.

6. Statistical Policy Office, Office of Information and Regulatory Affairs. Quality in Establishment Surveys. Office of Management and Budget. Statistical Policy Working Paper 15(39). July, 1988.