WEIGHTING AND ESTIMATION PROCEDURES FOR THE 1994 NEHIS

Leslie Wallace, Westat, Edward C. Bryant, Westat, David W. Chapman, Klemm Analysis Group,
David A. Marker, Westat, Christopher L. Moriarity, NCHS
Leslie Wallace, Westat, Inc., 1650 Research Blvd., Rockville, MD 20850

KEY WORDS: Nonresponse adjustment, trimming, post-stratification, self-employed, public sector, private sector

1. Introduction

The National Employer Health Insurance Survey (NEHIS) was conducted in 1994 by Westat, Inc., under contract to the National Center for Health Statistics (NCHS). The purpose of the NEHIS was to collect information on the health care insurance that U.S. businesses and governments provide for their employees. The survey collected information from employers on the names and types of health insurance plans (if any) offered to their employees, enrollments in these plans, the characteristics of the plans, the money paid for claims in the preceding plan year, and other related data.

The target sample size for the 1994 NEHIS was about 37,000 interviews for private establishments (i.e., specific business locations) and about 3,000 interviews for government agencies, for a total of about 40,000 interviews. The sample design was a stratified random sample of establishments. Strata were defined by state and size class in terms of the number of employees in the establishment. In the private sector, the number of employees in the "firm" containing the establishment was also used as a stratifier. For the public sector, type of government was included in the stratification process. In general, establishments in larger size classes were sampled at higher rates. An overview of the sample design is provided by Marker, et al. (1994). Additional details of the sample design, including the sources of the sampling frames, are provided by Westat (1994).

In order to minimize respondent burden, there was some subsampling of plans for establishments that offered a large number of plans. Because of the plan subsampling, it was necessary to develop two different sets of weights to use to compute survey estimates: one for the sample of establishments and another for the sample of plans. Although there was no subsampling required for about 95% of the responding plans, the subsampling required for the other 5% was sometimes complex, involving the subsampling of both establishments and plans from large companies that had substantial numbers of establishments selected for NEHIS. Some discussion of plan subsampling is given in Section 3, with more details being provided by Westat (1994).

This document provides a description of the weighting, estimation, and variance estimation methods used in NEHIS. The derivation of establishment and plan weights is described in Sections 2 and 3. The computation of survey estimates and variance estimates is described in Sections 4 and 5. As part of the estimation process, imputations are being made for selected variables, using the hot deck method. However, because of space limitations, a description of the specific imputation procedures used is not included here.

2. Establishment Weights

Establishment weights are needed to support analysis of establishment-level data, such as the number of establishments that offer insurance or the number of employees who have access to health maintenance organizations (HMOs). Establishment weights were computed in several steps: (a) computing base weights, (b) deriving nonresponse weight adjustments, (c) trimming excessive weights, and (d) post-stratification to independent universe counts (private sector only). The base weights (or basic sampling weights) were computed as the reciprocal of the establishment selection probability. These weights provide the basis for computing unbiased estimates of universe totals. In order to minimize the potential for nonresponse bias in survey estimates, nonresponse weight adjustments were then derived. Next, excessive weights were trimmed to reduce the impact of these weights on the variances of survey estimates. Finally, for the private sector, weights were post-stratified to align weighted establishment counts with adjusted Bureau of Labor Statistics (BLS) counts. Although weights for the public sector were not post-stratified, weighted government counts were compared to published estimates as a check on the weights.

These four weighting steps are described in more detail in Subsections 2.1 - 2.4. The final subsection, 2.5, describes the weighting of self-employed respondents with no employees.

2.1 Base Weights

Every establishment and government on the NEHIS sampling frames had a known, non-zero probability of selection, which was equal to the stratum sampling rate.

Except for the special situation discussed below, the base weights were equal to the reciprocal of the stratum sampling rate. Since the stratum sampling rates varied considerably, so did the base weights. In general, larger governments and establishments were selected with high probabilities, and therefore received lower base weights.

The base weights for non-locatable sample establishments were adjusted because of the uncertainty of their eligibility. To be specific, about 9,700 sample establishments were never located during the telephone interview process. Typically, no one answered at the telephone number after repeated attempts, and other resources (such as directory assistance) failed to turn up additional working numbers. Three percent of these cases were assumed to be eligible for NEHIS based on the observed eligibility rate in a sample of 50 of these cases located in Maryland. Ideally, 3 percent of the cases would be coded as eligible and 97 percent as ineligible for weighting purposes. However, individual cases could not be identified as eligible or ineligible. Thus the base weights of all the cases were adjusted by multiplying by 0.03 so that the entire group would represent the number of eligible establishments thought to be in the population. The largest base weights were 714.8 in the private sector and 177.2 in the public sector.

2.2 Nonresponse Adjustment

Several stages of nonresponse adjustment were required. In the first stage, establishments whose eligibility status had been determined were adjusted to account for establishments with unknown eligibility. The latter group differs from the non-locatable cases that received the 0.03 adjustment in that the phone number was verified but only minimal data were subsequently obtained. For the private sector, adjustments were calculated nationally within 22 cells. To form the cells, eligibility rates were calculated among the cases with known eligibility by state, firm size, and establishment size. Establishments in states with similar eligibility rates were grouped together for the adjustment. Likewise, since they had lower eligibility rates, smaller establishments were grouped separately from larger ones. For the public sector, adjustments were done separately by government type (counties, municipalities, special districts, school districts). The adjustment factor was the ratio of the sum of the weights of all sample cases (respondents, nonrespondents with unknown eligibility, eligible nonrespondents, and ineligible cases) to the sum of the weights of all sample cases except nonrespondents with unknown eligibility. The largest adjustment factors for stage 1 were 1.6 in the private sector and 1.2 in the public sector.

The remaining stages of nonresponse were done separately by state. The second stage adjusted for nonresponding establishments whose insurance status (whether they offered insurance) was unknown, and the third stage adjusted for nonresponding establishments whose insurance status was known. In the third stage, whether insurance was offered was used to form adjustment cells. The adjustments were performed separately by whether insurance was offered because this characteristic was correlated with both the response rate and the survey responses. For both of these stages, the adjustment factor was the ratio of the sum of the weights of eligible cases to the sum of the weights of the respondents. The largest adjustment factors for stage 2 were 1.3 in the private sector and 1.01 in the public sector. For stage 3, they were 1.6 and 1.5 in the private and public sectors, respectively.

Stage 4 was used to adjust for nonresponding certainties in the public sector. Unlike the other stages, the adjustment factor at stage 4 was calculated in terms of the number of employees at the establishment in order to better account for large establishments. This procedure was used only for the public sector because the largest governments were much larger than the largest establishments and more governments than private establishments were selected with certainty. The largest adjustment factor was 3.4 at this stage.

The initial adjustment cells for stages 2 and 4 were basically the sampling strata, formed by crossing state, firm size (1-49, 50-999, >999), and establishment size (unknown, 1 employee and no other locations, 1-5, 6-24, 25-49, 50-249, 250-999, >999) for the private sector; and state, government type, and establishment size (0, 1-5, 6-49, 50-249, 250-999, 1000-4999, 5000-9999, >9999) for the public sector. These cells were collapsed as described below. The collapsed cells from stage 2 were further subdivided by whether insurance was offered to form the initial cells for stage 3.

An attempt was made to avoid large adjustment factors and small adjustment cells (in terms of the number of establishments receiving the adjustment factor). This procedure was intended to reduce variances while increasing the potential for bias such that the overall mean square errors were reduced. In stages 2 and 3, initial adjustment cells were collapsed if the overall adjustment factor to that point was larger than 2, or the number of establishments receiving the adjustment was less than 6 and the adjustment factor was greater than 1. Note the focus was on the overall adjustment factor instead of the factor at each stage. In stage 4 the minimum cell size allowed was 2

governments, due to the fact that the adjustments were based on number of employees rather than number of governments.

Priorities were set for collapsing in an attempt to produce adjustment cells that would be homogeneous with respect to the survey data. For instance, collapsing was never done across state at any stage, nor across whether insurance was offered at stage 3. For the private sector, the first choice was to collapse across establishment size within the same firm size, using smaller size categories first (to avoid collapsing noncertainties with certainties). For the public sector, government types were combined (except school districts) before government size. The cell size and adjustment-factor size requirements were met using these guidelines in all but a few cells. In these cases, the rules were relaxed (for example, crossing firm size in the private sector) when the factors otherwise were unusually large. The largest overall adjustment factors were 2.02 in the private sector and 4.7 in the public sector.

2.3 Weight Trimming

The nonresponse-adjusted weights were reviewed for possible trimming. Trimming was needed primarily because some measures of size used to assign establishments to sampling strata were inaccurate. Establishments were placed into sampling strata based on the number of employees given in the sampling frame but will be analyzed based on their reported number of employees. Small establishments were sampled at lower rates than large ones. If an establishment was actually much larger than the frame indicated, the establishment weight would be much larger than the weights of other establishments of the same size. Conceivably, a few establishments could dominate certain subgroup estimates.

For the private sector, establishment weights were trimmed if the weighted difference in the establishment's size (frame size versus reported size) accounted for at least 8 percent of the estimated number of employees in firms of similar size in the same state. For the public sector, weights were trimmed if the weighted difference in the establishment's size accounted for at least 5 percent of the estimated number of employees in the state. Another criterion for trimming was that the reported establishment size had to be at least ten times larger than the frame size. Weights were trimmed for four governments and 127 private establishments. Each weight was trimmed to what it would have been if the establishment had been placed in a sampling stratum based on its reported size instead of its frame size. For trimmed establishments, adjustment factors ranged from 0.03 to 0.53 in the private sector and 0.08 to 0.15 in the public sector.

2.4 Post-stratification

For the private sector, the trimmed establishment weights were post-stratified to be consistent with BLS employee counts. Counts of employees were used instead of counts of establishments because a major focus of NEHIS is on data correlated with number of employees (cost data, number of employees with access to health care, etc.). Also, the definition of an employee was more consistent across surveys than the definition of an establishment. Furthermore, the number of employees is more stable over time than the number of establishments.

Post-stratification was done within 404 cells formed by crossing state, industry grouping (goods producing versus services), and reported establishment size (1-9, 10-49, 50-249, >249). The original total of 408 cells (50 states and the District of Columbia by 2 industry groupings by 4 establishment sizes) included 4 cells that were paired (collapsed) with others to meet minimum cell size or adjustment-factor size requirements.

No perfect data source was available to provide the control totals due to differences in reference time periods, frame coverage, and definitions. The control totals used for NEHIS were March 1994 universe counts of employees used for BLS's Employment and Earnings Survey. These figures were increased in 3 major ways to account for nearly 8 million people NEHIS had included as employees but BLS had not. The first addition was for Washington state corporate executives (estimated to be 37,700), who are not covered by employment insurance. The second addition was for over two million employees working in certain standard industrial classification (SIC) groups (mainly railroads and churches) not covered by employment insurance. The third addition was to account for self-employed persons with employees. While their employees had been counted, the self-employed people themselves were not included in the BLS figures. The post-stratification factors for each cell ranged from 0.4 to 2.1, and the national post-stratification factor was 0.99.

For the public sector, no source existed that was clearly better than the sampling frame to provide control totals for post-stratification. Instead, the public sector establishment weights were checked for reasonableness against BLS December 1993 Employment and Earnings data. The total number of government employees (federal, state, and local) and the number of state and local government employees were estimated for each state using the product of the final establishment weight and the reported

establishment size. These figures were compared to similar figures based on the Employment and Earnings Survey.

The NEHIS estimate of total government employees was within 10 percent of the Employment and Earnings figure for 16 states, and within 20 percent for 33 states. The national NEHIS estimate was 10 percent lower than the Employment and Earnings figure. For state and local governments, the NEHIS estimate was within 10 percent of the Employment and Earnings figure for 14 states, and within 20 percent for 28 states. The national NEHIS estimate was 12 percent lower than the Employment and Earnings figure. It is unclear whether the NEHIS estimates are low or the Employment and Earnings data are high. The latter are benchmarked to 1987 Census of Governments (COG) data, which may overstate employment due to recent downsizing. It would be interesting to repeat the exercise once the Employment and Earnings data are benchmarked to the 1992 COG.

2.5 SENE Weights

A sample of self-employed persons with no employees (SENEs) was selected and processed separately from the private and public sector surveys. Self-employed respondents from the last six months of the 1993 National Health Interview Survey (NHIS) constituted the SENE sample. Establishment weights were constructed for SENEs as the product of two components: a base weight and a nonresponse adjustment. No control totals were available for the number of self-employed people with no employees, so SENE weights were not post-stratified.

The base weight for SENEs was the final weight from NHIS, which included adjustments for nonresponse and post-stratification (to overall NHIS control totals, not restricted to SENEs). The base weights of the respondents and ineligible cases were adjusted to account for nonrespondents. Nonrespondents included people whose eligibility for NEHIS could not be determined and people who were eligible but did not respond. The adjustment factor was the ratio of the sum of the weights of all sampled cases (respondents, nonrespondents, and ineligible cases) to the sum of the weights of the respondents and ineligible cases. This procedure assumed that the eligibility rate for all nonrespondents was equal to the eligibility rate observed for respondents.

The adjustment was done within initial cells formed by crossing Census Region, type of primary sampling unit (PSU), NHIS sampling stratum, and PSU. The initial cells were collapsed to yield at least 6 completes in each cell with some nonresponse, and adjustment factors no larger than 2. The priority was to collapse PSU in the same sampling stratum. Only a few cells needed collapsing. The base weights ranged from 634 to 17,638, differing by a factor of 28. The largest nonresponse adjustment was 1.9, and the final weights differed by a factor of only 21.

3. Plan Weights

Plan weights are needed to support analysis of plan-level data such as premiums, deductibles, and copayments. "Plan-level" actually means plan within establishment (or plan-establishment pair) rather than the plan across all establishments where the plan is offered. Responding plans in responding establishments have final plan weights. The final plan weight is the product of four or five components. The first component is the final weight for the establishment from which the plan was selected. Next is the plan subsampling weight which is the inverse of the conditional probability of selecting the plan from the establishment, given that the establishment has been selected. Next are two nonresponse adjustments, which were still underway at the time of this writing. The nonresponse adjusted weights will be reviewed to see if trimming excessive weights is necessary. The weights will not be post-stratified since suitable control totals are not available.

Plan subsampling was used to reduce respondent burden for establishments that offered more than five health plans, and for firms with large numbers of establishments in the NEHIS sample. In these cases, up to 13 health plans were subsampled, with no more than 5 health plans subsampled from any one establishment. In the largest private firms, one establishment was subsampled from each of ten states, and then plans were selected within the subsampled establishments. A variety of plans was subsampled including major medical and single service company-wide plans and major medical local plans (i.e., not company-wide). No plan subsampling was necessary for 95 percent of the plans in both the public and private sectors. The largest plan subsampling weights were 114.0 in the private sector and 14.5 in the public sector.

Plan nonresponse adjustments will be done in two stages. In stage 1, plans that were either self or fully insured will be adjusted to account for nonresponding plans whose self or fully insured status was unknown. In stage 2, responding self insured plans will be adjusted to account for nonresponding self insured plans, and responding fully insured plans are being adjusted to account for nonresponding fully insured plans. Plan nonresponse is being done in two stages because whether the plan was fully or self insured is thought to be highly correlated with the survey responses. The questionnaire was considerably different

for these two situations. It was preferable to use self or fully insured status in forming adjustment cells, but this status was not known for most nonresponding plans. Thus plans for which self or fully insured status was not known were handled separately than plans for which self or fully insured status was known.

The adjustment factors for both stages will be functions of plan enrollment. The initial cells in stage 1 will be formed by crossing the cell from stage 2 of establishment nonresponse with plan type (conventional, preferred provider, point of service or HMO, dental, other single service). These cells will be collapsed in a manner similar to that used for establishment weights. The collapsed cells from stage 1 will be subdivided by whether the plans were self or fully insured to form the initial cells for stage 2.

As with establishment weights, the aim is to limit the maximum adjustment factor (to 2 at stage 1 and overall) and the minimum cell size allowed (to 6 for the private sector and 2 for the public sector) during plan nonresponse. The minimum cell size has been reduced from 6 for establishment nonresponse to 2 for plan nonresponse so that plan-level variables can be used to form adjustment cells in addition to the establishment-level ones. The guidelines for collapsing the initial cells are the same as those used for the establishment weights, with the following enhancements. Collapsing self or fully insured plans together in stage 2 is not allowed. In both stages, after trying to collapse across establishment size, dental is being collapsed with other single service plans. Collapsing big establishments (at least 250 employees) with small ones (fewer than 250 employees) is being avoided. Meeting the minimum cell size and maximum adjustment factor requirements using these guidelines is difficult, but the vast majority of cells with nonresponse will still meet the requirements.

Plans were not subsampled for SENEs since the number of plans offered at any one SENE establishment was expected to be small. Plan nonresponse was negligible. Thus, the plan weights for SENEs are identical to the establishment weights.

4. Survey Estimates

Most of the estimates that will be generated from the 1994 NEHIS results will be totals, percents, or means. In the case of percents and means, these will be computed as a ratio of estimated totals. Therefore, the focus of this section is on estimates of survey totals.

Two basic types of estimates will be computed: those at the establishment level, which will be based on establishment weights, and those at the plan level, which will be based on plan weights. An example of an establishment level estimate is the estimated number (or percent) of establishments that offer employer-sponsored health insurance to their employees. An example of a plan level estimate is the estimated number of employees enrolled in a health insurance plan with family coverage.

Estimates for establishment level characteristics are described in Section 4.1, while estimates for plan level characteristics are discussed in Section 4.2.

4.1 Establishment Level Estimates

An estimate of an establishment total, Y, will be computed as the weighted sum of the y-values from the respondent establishments:

$$\hat{Y} = \sum_{h=1}^{L} \sum_{i=1}^{n_h} w_{hi} * y_{hi} \qquad (1)$$

where

w_{hi} = the final weight assigned to the i^{th} establishment in stratum h, as defined in Section 2,

y_{hi} = the value of the Y variable for the i^{th} establishment in stratum h.

The estimated establishment totals computed from the NEHIS results will be of two general types: simple establishment counts, and totals such as employee totals and total health insurance costs. For estimates of establishment counts, the y_{hi} variable in Equation (1) is a 0-1 variable, which takes on the value 1 for each respondent that has the characteristic being estimated, and 0 otherwise. To compute an estimated establishment proportion, the estimated total in Equation (1) is divided by the sum of the weights of the respondents.

Totals such as employee totals will also be estimated from Equation (1). An example would be the estimated number of employees eligible for employer-sponsored health insurance. In this case, the y_{hi} variable in Equation (1) is the corresponding employee count (e.g., number of employees eligible for health insurance) for the i^{th} establishment in stratum h. To estimate the proportion of employees eligible for health insurance, the estimated total from Equation (1) would be divided by the estimated total number of employees.

Estimates for population domains will be computed from Equation (1) by limiting the sum to the members of the domain, or by adding a 0-1 (indicator) variable to the Equation. The indicator variable would take on

the value 1 for members of the domain of interest, and 0 otherwise.

4.2 Plan Level Estimates

Plan level estimates will consist of plan enrollment estimates, means per enrolled employee, and percents of enrolled employees that have some characteristic. An example of an enrollment total is the estimated total number of employees enrolled in an HMO. An example of a percent would be the percent of enrolled employees that are in plans that cover childhood immunizations. An example of a mean would be the mean premium paid by employees enrolled with single coverage. Each of these estimates is either a total or a ratio of two estimated totals.

Estimates of plan level totals will be computed as weighted totals of appropriate plan enrollments, using an equation similar to Equation (1) for establishment level estimates, with a third subscript added for the plan within the establishment:

$$\hat{Y} = \sum_{h=1}^{L} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} * y_{hij} \qquad (2)$$

where

w_{hij} = the final weight assigned to the j^{th} plan enumerated in the i^{th} establishment in stratum h, as defined in Section 3,

y_{hij} = the value of the Y variable (e.g., plan enrollment) for the j^{th} plan enumerated in the i^{th} establishment in stratum h.

It is recommended that plan level estimates be based on enrolled employees, rather than on a universe of plans. This is because the plan universe being sampled is the collection of all establishment-plan combinations, which would include many individual plans several times. Therefore, if a statistic like the average number of persons enrolled per plan were computed, it would have to be clearly stated that this average is over all establishment-plan pairs.

For domain estimates, Equation (2) will be used with the sums being taken over domain members, or with a 0-1 indicator variable added to the equation. As before, the indicator variable would take on the value 1 for plans in the domain of interest (e.g., HMO plans), and take on the value 0 otherwise.

5. Variance Estimation

Plans for variance estimation for the 1994 NEHIS are to produce design-based sampling errors that take account of the complex sample design (stratification for the establishment sample, and clustering for the plan subsample). SUDAAN software will be used to produce most direct estimates. SUDAAN uses a Taylor series linearization method for variance estimation. Replication software such as WESVAR and/or VPLX will be used to produce some sampling error estimates, to provide a consistency check with SUDAAN's results.

The basic strategy for variance estimation is the same for both establishment-level variance estimates and plan-level variance estimates. The sample establishment or government will be treated as the "ultimate cluster" (Hansen, Hurwitz, and Madow (1953), page 242). As described previously, establishments were the primary sampling units and plan sampling usually was nested within the establishment; however, the sampling for multi-establishment firms did not necessarily follow this pattern. In the absence of a variance estimation technique that is specifically designed for this situation, nesting of plans within establishments for subsampling will be assumed for variance estimation purposes.

Although plans are to compute some direct estimates of sampling error, it is not feasible to compute a direct estimate for every possible statistic that could be produced from NEHIS data. For this reason generalized variance models will be produced. Groups of NEHIS statistics will be based on similar statistic characteristics (e.g., plan-level statistics of plan enrollment) and goodness of model fit. A relatively small number of models will be developed that fit the observed data well.

References

Hansen, M., Hurwitz, W., and Madow, W. (1953). Sample Survey Methods and Theory, Volume I, Wiley.

Marker, D., Bryant, E., and Moriarity, C. (1994). National Employer Health Insurance Survey Sample Design, ASA 1994 Proceedings of the Section on Survey Research Methods, Volume I., pp.264-269.

Westat, Inc. (1994). National Employer Health Insurance Survey Design, a project report submitted to the National Center for Health Statistics, May 5, 1994.