

PATTERNS OF AND ADJUSTMENTS FOR NONRESPONSE IN AN NHIS SUPPLEMENT

Steven L. Botman

NCHS, 6525 Belcrest Rd., Hyattsville, MD 20782

KEY WORDS: Survey participation, weighting

INTRODUCTION AND SUMMARY

Detailed information on nonrespondents is often not available in household surveys, but is available when a household sample is drawn from a larger survey. This paper explores nonresponse patterns for a National Health Interview Survey (NHIS) supplement, where respondents to the NHIS basic health and demographic questionnaire ("the core") are taken as the larger survey. This paper addresses the supplements that are administered to a randomly selected adult in each family in the NHIS, with proxy response not accepted; there are other designs for NHIS supplements. A number of variables may be related to nonresponse in face-to-face household surveys. Botman and Thornberry (1993) present completion rates for an NHIS supplement similar to that in this paper by several variables, including age, sex, and race. More generally, Groves (1989) notes some correlates of nonresponse in household interview surveys. This paper presents nonresponse patterns for such an NHIS "sample adult" supplement based on the response to the larger survey.

This paper presents data generally showing narrow variation in levels of NHIS supplement participation using the information collected in the larger survey. Clearly the choice of other variables or other response categories for these variables, however, may affect the patterns.

Prior to survey adjustment for nonresponse, for an estimated mean in a domain of study, nonresponse bias is a function of the nonresponse rate and the difference in the means between the respondents and nonrespondents in that domain. For a survey where the overall response rate is high and narrow variation occurs in response rates, nonresponse bias is not a major part of the survey error structure. This paper provides some confirmation of this in the case of the NHIS supplement.

In the design of the sampling weights for an NHIS supplement, supplement nonresponse is typically taken into account in poststratification. The poststratification employed in the NHIS and the supplements is by age, sex, and race. That is, no discrete weighting adjustment is employed for supplement nonresponse. In an alternatively designed set of experimental sampling weights for a supplement, to explore this matter, we compute and use a set of experimental weights with both a discrete supplement nonresponse

weighting adjustment and a revised poststratified weighting adjustment. The cells for the nonresponse adjustment are determined by a Chi-squared AID (CHAID) analysis done using SI-CHAID, published by Statistical Innovations Inc.

For a national percentage estimate of adult smoking status, the revised post-stratified survey estimate based on an estimator with a discrete adjustment for supplement nonresponse did not differ from that based on the usual supplement estimator, which had no explicit survey adjustment for supplement nonresponse. Clearly for subdomains or means for other characteristics of interest, this result may not hold. Still it provides evidence that a discrete adjustment for nonresponse is not always needed.

NHIS BACKGROUND

The NHIS is a large, continuous face-to-face household survey conducted by the National Center for Health Statistics (NCHS). The Bureau of the Census under a contractual arrangement participates in the planning of the NHIS and is the NHIS data collection agent. While the NHIS sampling plan implemented in 1985 will be used through 1994, during each survey year the NHIS data collection plans change.

The NHIS basic health and demographic questionnaire consists of a fixed set of health and sociodemographic questions--this questionnaire usually remains unchanged for about 10 years. This portion of the survey is used, for example, to develop annual estimates of acute and chronic conditions, hospital stays, medical visits, and limitations of activities. NHIS core data are collected family style on all members of the household. Self-response is used for all adults at home at the time of the interview and proxy response for children and for adults not available or unable to respond for themselves. This survey component averages about 30 minutes.

The NHIS supplement questionnaires address health issues or topics identified as of particular interest within the U. S. Public Health Service. These topics are usually included in the NHIS for an entire survey year. Data collection for the core and supplement questionnaires together typically requires about 90 minutes for data collection.

For the basic health and demographic questionnaire, NCHS has obtained response rates usually about 95 percent. For the supplements based on a randomly selected adult in each family in the NHIS (with proxy not accepted), NCHS typically obtains completion rates

from 85-90 percent (relative to response to the NHIS basic health and demographic questionnaire).

METHODS

Prerequisites. In order to conduct the analysis in this paper, some processing is needed. Such analysis is based on being able to determine the cases designated for supplement sample and their disposition. The NHIS supplementary data sets generally do not include a record for each nonrespondent. Because of this, we have to infer those cases designated for the supplement sample.

Operationally in each responding NHIS sample household the interviewer enumerates the adults eligible for such a sample person supplement. These adults are listed in a prespecified order on numbered lines on the face sheet of the supplementary questionnaire. Each line identifies one adult. Based on the number of lines used, the interviewer refers to the label for the case and determines the number of the individual to be selected for the sample and then attempts to administer the supplement questionnaire.

In a relatively small proportion of cases, the keyed number for the sample person is not consistent with the number for the sample person keyed from the label. The interviewer could have incorrectly selected the sample line number; or the interviewer could have selected the sample line number correctly but an error could have been made in survey processing. We did not use these cases. The figures on supplement participation may thus not agree those presented elsewhere.

Tabulation and Analysis. Using information from the basic NHIS basic health and demographic questionnaire, we tabulated the supplement completion rate by several NHIS core variables.

A CHAID analysis was used to segment the sample by supplement participation. In a sense the CHAID segments subgroups with maximum statistical differences in survey completion rates. As part of the segmentation process we identified a number of variables that might be related to the kind of data collected in the supplement.

In the CHAID analysis some variables were not used. For example, the month of respondent birth was not used in this segmentation. In fact some variation did occur by month of birth. Some of this variation could be explained. When a respondent did not provide a month of birth, the same value was imputed. Clustered in this month were a disproportionate number of persons who were not good survey respondents.

RESULTS

Table 1 shows little variation in supplement

completion over a variety of respondent characteristics. There are some factors that do have some larger differences, especially when the response is unknown for a nonthreatening data request (e.g., educational attainment or health status). Some of these persons may be tenuous members of the household. For persons with a telephone number but without a reported telephone number in the NHIS, the situation may be operational (the interviewer could not check back with the household by telephone to schedule a productive callback or could reflect some respondent hostility to the interview), or may be an interview break off. It's not obvious that the true mean for domains of persons with unknowns for nonthreatening items would differ from that for the residual population.

Figure 1 shows a CHAID analysis on response to the supplement. Many results are expected. The completion rates are highest among households of smaller size. For a family of size 1, it's easier to contact the designated sample person once the household is actually contacted; it, however, is more difficult to contact such a household though. With a household of larger size the sample adult selected for the supplement more often may be absent when the household is initially contacted for the survey.

Supplement completion rates tended to be slightly lower in the largest metropolitan areas. Again the rates have been noted by Groves and others to be lower in metropolitan areas, but the participation pattern is generally not noted as affected by the particular size of MSA. Some smaller MSAs have high response rates. Botman and Thornberry (1992) identified some variation in NHIS response rates for particular large MSAs.

As noted, when Figure 1 was used to form weighting cells for a supplement nonresponse adjustment, the revised poststratified survey estimator yielded the same national smoking level estimate for population aged 18 and older as did the usual estimator. While these estimators may have yielded different figures for either other characteristics or this characteristic for other domains, for a key national estimate they yielded the same figure. At least in this exploratory analysis for the full population a discrete nonresponse adjustment was not helpful. Use of a discrete nonresponse adjustment would have also required dropping from the survey of file of some cases with questionably correct implementation of the sampling rule.

REFERENCES

Groves, R. M., (1989) *Survey Errors and Survey Costs*, John Wiley.

Botman and Thornberry (1992), Subnational Survey Design Feature Correlates of Nonresponse, Presentation at 1992 APHA meeting.

Botman and Thornberry (1993), Survey Design Features Correlates of Nonresponse, 1992 ASA Proceedings.

Table 1. 1991 HPDP Supplement Disposition Status by NHIS Basic Health and Demographic Variables

Basic Health and Demographic Variables	DISPOSITION		
	ALL	NON RESPONDING	RESPONDING
	N	PERCENT	PERCENT
ALL	47,424	10	90
METRO STATUS			
MSA IN CENTER CITY	15,991	10	90
MSA NOT IN CENTER CITY	20,578	9	91
NONMSA AREA	10,855	9	91
CENSUS REGION			
EAST	9,761	9	91
NORTH CENTRAL	11,841	8	92
SOUTH	15,815	11	89
WEST	10,007	9	91
MSA POPULATION SIZE			
MSA 1,000,000 OR MORE	20,289	11	89
MSA 200,000--999,999	12,528	9	91
MSA 100,000-199,000	2,981	8	92
MSA UNDER 100,000	771	6	94
NON-MSA, OTHER URBAN	4,462	8	92
NONMSA, RURAL	6,393	10	90
AGE			
18-24	5,375	10	90
25-44	20,823	10	90
45-64	11,941	10	90
65-69	2,784	7	93
70-74	2,524	7	93
75+	3,977	10	90
SEX			
MALE	20,527	12	88
FEMALE	26,897	7	93

Table 1. 1991 HPDP Supplement Disposition Status by NHIS Basic Health and Demographic Variables, continued

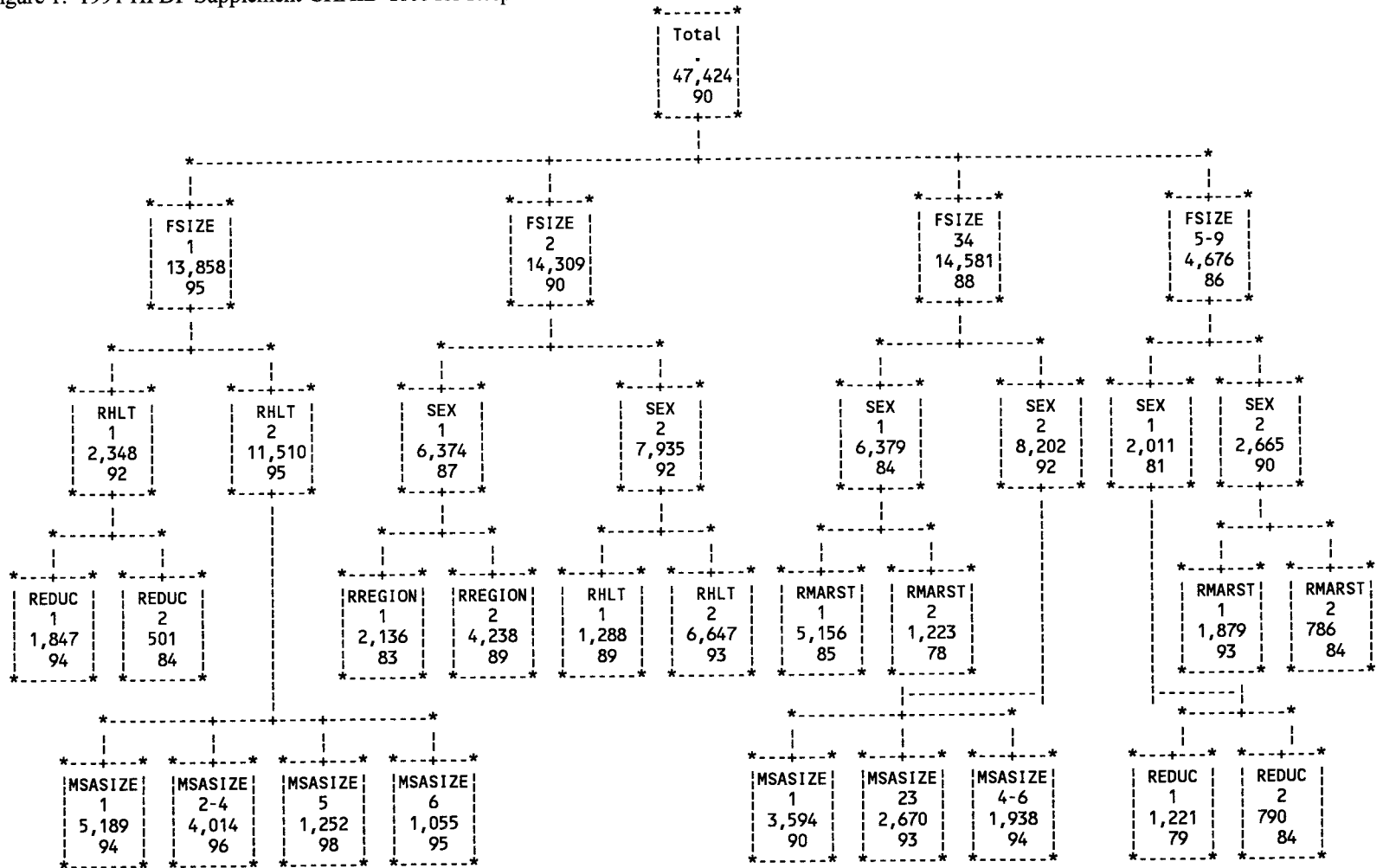
Basic Health and Demographic Variables	DISPOSITION		
	ALL	NON RESPONDING	RESPONDING
	N	PERCENT	PERCENT
RACE			
WHITE	39,276	9	91
BLACK	6,630	11	89
OTHER	1,518	12	88
HISPANIC STATUS			
HISPANIC	3,411	10	90
NONHISPANIC	44,013	9	91
MARITAL STATUS			
MARRIED, SPOUSE IN HHD	25,262	11	89
MARRIED, NO SPOUSE IN HHD	566	8	92
WIDOWED	5,266	7	93
DIVORCED	5,023	5	95
SEPARATED	1,505	5	95
NEVER MARRIED	9,508	9	91
UN	294	83	17
FAMILY SIZE			
1	13,858	5	95
2	14,309	10	90
3-4	14,581	12	88
5 OR MORE	4,676	14	86
EDUCATION ATTAINMENT			
LESS THAN HIGH SCHOOL	10,152	10	90
HIGH SCHOOL	17,432	9	91
SOME COLLEGE	9,844	8	92
COLLEGE OR HIGHER	9,522	8	92
UN	474	77	23
DETAILED FAMILY INCOME			
UNDER 10,000	6,469	4	96
10,000-19,999	8,156	6	94
20,000-29,999	6,820	6	94
30,000-39,999	5,469	7	93
40,000-49,999	4,040	8	92
50,000 AND OVER	7,585	9	91
UN	8,885	22	78

Table 1. 1991 HPDP Supplement Disposition Status by NHIS Basic Health and Demographic Variables, continued

Basic Health and Demographic Variables	DISPOSITION		
	ALL	NON RESPONDING	RESPONDING
	N	PERCENT	PERCENT
FAMILY INCOME STATUS			
LT 20,000	18,132	7	93
GE 20,000	27,595	9	91
UN	1,697	44	56
MAJOR ACTIVITY			
WORKING	28,013	10	90
KEEPING HOUSE	9,500	6	94
GOING TO SCHOOL	2,641	9	91
SOMETHING ELSE	6,934	12	88
UN	336	17	83
HEALTH STATUS			
EXCELLENT	15,476	10	90
VERG GOOD	13,579	8	92
GOOD	11,848	9	91
FAIR	4,510	9	91
POOR	1,846	15	85
UN	165	44	56
EMPLOYMENT STATUS			
IN LABOR FORCE	29,007	10	90
UNEMPLOYED	1,563	7	93
NOT IN LABOR FORCE	16,854	9	91
NHIS/HOUSEHOLD TELEPHONE STATUS			
TELEPHONE HOUSEHOLD AND # REPORTED	42,462	9	91
TELEPHONE HOUSEHOLD AND # NOT REPORTED	2,012	25	75
NONTELEPHONE HOUSEHOLD	2,767	13	87
UN	183	23	77

Figure 1: 1991 HPDP Supplement-CHAID Tree for Response

938



Note: variables and response categories in CHAID tree

FSIZE = Family Size; RHLT: 1 = Fair, poor, or UN Health Status, 2 = other; SEX: 1 = Male, 2 = Female;
 RREGION: 1 = South, 2 = Other Census regions; RMARST: 1 = Married, 2 = Other; MSASIZE: 1 = MSA 1,000,000 or more,
 2 = MSA 250,000 - 999,999, 3 = MSA 100,000 - 249,000, 4 = MSA under 100,000, 5 = nonMSA other urban areas,
 6 = NonMSA rural areas.