AN INDICATION OF THE EFFECT OF NONINTERVIEW ADJUSTMENT AND POST-STRATIFICATION ON ESTIMATES FROM A SAMPLE SURVEY

Martha J. Banks, University of Chicago

This paper is based on research done at the Center for Health Administration Studies, University of Chicago. Using data from a national sample survey of medical care use in 1970, we investigated various components of total survey error and methods to improve the validity and reliability of the survey estimates. The results of this study will appear in the upcoming book, Total Survey Error: Bias and Random Error in Health Survey Estimates, edited by Ronald Andersen, Judith Kasper, and Martin Frankel. The data were collected and processed by the National Opinion Research Center. The funding for this methodological investigation, as well as for the data collection and basic analysis, was provided by the National Center for Health Services Research. The National Center for Health Statistics also provided valuable support.

This paper concentrates on two features of this investigation, on adjustments for nonresponse and on post-stratification adjustment. Both are relatively easy to implement and so could be used in situations in which other data adjustment techniques might not be felt to be worthwhile.

The basic rationale for nonresponse adjustment might be described as follows: In almost any survey there will be cases which were designated for interview but which were not actually interviewed. Some potential respondents may have refused to be interviewed; others were not at home when repeated interview attempts were made. Making no adjustment for nonresponse implicity assumes that nonrespondents do not tend to differ from respondents in any characteristic of interest. The degree to which they do differ is proportional to the amount of bias introduced by ignoring the nonresponse problem.

Any approach to nonresponse adjustment consists of two elements. First, the population must be categorized into subgroups and the response rate for each group must be determined. The categories chosen to form the subgroups should not only be correlated with characteristics of interest in the study, but be able to be determined without having to obtain the information from the potential respondents themselves.

The second major element in nonresponse adjustment is that of determining the values to impute to the nonrespondents. It usually is reasonable to assume that respondents and nonrespondents falling into the same category tend to have the same characteristics. Sometimes however, we have evidence that respondents and nonrespondents in the same category have measurably different characteristics. This evidence may come from the current study or from external sources. The external data might be from a previous study which either had subsampled nonrespondents or had access to the administrative records of both survey respondents and nonrespondents.

With the data we had available, we chose to try two alternative nonresponse adjustment procedures. Both assume that respondents and nonrespondents within the same category tend to have the same characteristics. Thus the overall results of the two methods differ only because of different choices of categories. The first method creates categories based on geographic location. The second uses information about the reason that no interview was obtained, whether refusal or never-at-home.

Before discussing each method, I first need to discuss the sample used in this investigation. In early 1971, persons in 3880 households were interviewed about the use and cost of health services used during 1970. In all, data were collected for 11,619 individuals. The sample was an area probability sample of the noninstitutionalized population of the continental United States. The sampling procedures oversampled poor persons living in the inner city, persons 65 and over, and rural residents. Naturally, weighting was used to adjust for this oversampling. The weighted nonresponse rate in the survey was 18 percent.

The geographic nonresponse adjustment method used primary sampling unit and sub-sample as the category determinants. This effectively grouped cases within PSUs by the presence of the poor and/or the elderly.

The other method of noninterview adjustment used categories based upon the reason that no interview was completed. To adjust for cases that refused to be interviewed, we increased the weights of those respondents who were not completely cooperative, breaking appointments with the interviewer and so on. The majority of nonrespondents in this study were refusals. To adjust for other types of nonresponses, those due to never being able to find anyone at home, we increased the weights of completed cases according to the number of calls needed to complete the interviews.

Table 1 compares the percentage of interviewed households that received various levels of noninterview adjustment weights according to each method. Most of the sample households were given weights between 1.02 and 2.00 by the geographic method, while most received a weight outside this range from the adjustment based on the reason that no interview was obtained.

TABLE 1 Distribution of sample households by noninterview adjustment factor

INFORMATION USED IN NONINTERVIEW ADJUSTMENT	NONINTERVIEW ADJUSTMENT FACTOR						TOTAL
	1.00	Over 1.00 thru 1.02	Over 1.02 thru 1.15	Over 1.15 thru 1.30	Over 1.30 thru 2.00	Over 2.00 thru 2.40	
Geographic	15.8	0.0	24.8	34.1	24.3	.1	100.0
Reason no interview was completed	67.3	12.7	3.4	1.0	2.9	12.8	100.0

TABLE 2 Effect of nonresponse adjustment on the distribution of sample persons, on estimates of mean number of physician visits for persons seeing a physician, and on estimates of mean hospital expenditure per admission

CHARACTERISTIC	PERCENT OF WEIGHTED SAMPLE PERSONS				MEAN PHYSICIAN VISITS PER PERSON				MEAN EXPENDITURE PER ADMISSION			
		Adjusted for Nonresponse				Adjusted for Nonresponse				Adjusted for Nonresponse		
	Unadjusted	Adjusted Using External Data	Adjusted Using Geographic Information	Adjusted Using Reason No Interview Obtained	Unadjusted	Adjusted Using External Data	Adjusted Using Geographic Information	Adjusted Using Reason No Interview Obtained	Unadjusted	Adjusted Using External Data	Adjusted Using Geographic Information	Adjusted Using Reason No Interview Obtained
Demographic												
Age of oldest family member												
Less than 65 years	86.2%	86.5%	86.2%	86.3%	5.4	5.5	5.4	5.3	$640	$662	$642	$624
65 years or more	13.8	13.4	13.8	13.7	7.8	7.9	7.8	7.7	863	886	877	857
Family income												
Nonpoor	77.0	77.3	77.1	77.8	5.5	5.6	5.6	5.5	674	697	679	643
Poor	23.0	22.8	22.9	22.2	6.5	6.6	6.6	6.5	715	737	717	749
Race												
White	87.9	87.9	88.0	88.3	5.7	5.7	5.7	5.6	684	709	691	659
Nonwhite	12.1	12.1	12.0	11.7	6.0	6.1	5.9	5.9	688	711	669	770
Residence												
Rural nonfarm	24.5	23.7	24.4	24.4	5.4	5.4	5.4	5.3	557	573	565	534
Rural farm	6.8	6.2	6.8	6.6	5.6	5.6	5.6	5.7	575	586	562	574
SMSA central city	29.8	31.2	29.4	29.7	6.0	6.1	6.1	6.0	727	757	745	735
SMSA other urban	26.9	27.0	27.2	26.5	5.6	5.6	5.6	5.4	910	941	904	904
Urban nonSMSA	12.1	12.0	12.2	12.8	6.1	6.2	6.2	6.1	444	458	446	431
Perceived and Evaluated Health												
Perception of health												
Excellent	37.8	-[a]	37.9	38.1	3.6	-[a]	3.6	3.5	488	-[a]	495	475
Good	43.0	-	42.7	42.9	5.3	-	5.3	5.2	609	-	619	614
Fair	12.4	-	12.4	12.1	9.1	-	9.2	8.9	698	-	712	682
Poor	3.9	-	3.9	3.9	14.2	-	14.2	14.5	868	-	877	862
Number of diagnoses												
One	28.2	-	28.2	28.5	3.8	-	3.8	3.8	626	-	643	617
Two	17.6	-	17.7	18.2	5.6	-	5.7	5.8	573	-	578	620
Three	9.2	-	9.1	9.0	8.0	-	8.0	7.8	563	-	559	544
Four or more	9.0	-	9.0	8.4	11.6	-	11.7	11.6	883	-	886	839
Total	100.0%	100.0%	100.0%	100.0%	5.7	5.8	5.8	5.7	$685	$707	$689	$670

[a]Data necessary to provide these estimates are unavailable.

Results from each nonresponse adjustment method were compared with each other and with data unadjusted for nonresponse. These appear in Table 2. Additional columns in this table are labeled "adjusted using external data." The limited number of estimates given in these columns were obtained by using data from various other health surveys to estimated differences between respondents and nonrespondents.

While discussing the data, I would like to stress that the table contains the results we have obtained from using each method of noninterview adjustment. I do not want to make any predictions about what the results would be expected to be if these methods were applied to a number of similar data sets.

Table 2 provides information about the effect of noninterview adjustment on the distribution of sample persons, on the mean number of physician visits for persons with visits, and on the mean hospital expenditure per admission. Most differences in the table are very small. However, all three sections shown the adjustment based on geographic information had less effect on the means than did the adjustment based on the reason that no interview was obtained. The noninterview adjustment seems to have had more effect on hospital expenditures than on physician visits, at least for totals and among the demographic characteristics.

None of this discussion has attempted to suggest which type of adjustment produces the most accurate estimates or even whether or not the time spent doing any type of adjustment for nonresponse is time well spent. In fact in most data collection there is no way to find out what would be the response of all nonrespondents. Therefore there is no way to determine the improvement in the estimates caused by noninterview adjustment. We can only measure the change it makes in the unadjusted estimates.

In our examples few of the estimates adjusted for nonresponse are very different from the unadjusted estimates. However, a well thought-out plan for noninterview adjustment usually is worth making, since doing so is fairly simple. Further, the benefits of noninterview adjustment probably are increasing, since the response rates of most surveys have been declining for at least a decade.

A fairly firm plan for noninterview adjustment should be devised before the study interviews are conducted, so that the desired information used in forming noninterview categories can be collected both for the respondents and for the nonrespondents.

As previously stated, noninterview adjustment also requires values to impute to the nonrespondents in each category. Unless there is firm evidence to the contrary, it would seem best to assume that respondents and nonrespondents falling into the same category are otherwise identical. Doing so usually is preferable to using external data to estimate values to impute to nonrespondents. Definitional and procedural differences between the current data and the external data require caution in adapting the results from the external data sources. It seems unlikely that the cost and time spent locating and adapting external data could often be justified.

It is difficult to choose between the two adjustment methods which use internal data, given the limited information available on the effect of each. I feel that the use of either is somewhat preferable to performing no noninterview adjustment at all, but either set of categories could be used.

Noninterview adjustment is a relatively inexpensive method of reducing nonresponse bias somewhat, but it certainly is no substitute for obtaining actual responses from as many designated respondents as possible. Adjustment should not be used as an excuse for a high nonresponse rate!

The reasons for post-stratification adjustment can be summarized as follows: Compared to the original population, a sample chosen from that population will exhibit chance differences in nearly all possible variables. Usually there are a number of characteristics which are correlated with the dependent variables of interest in the study and for which more reliable estimates exist. Thus the study data generally can be improved by applying a set of factors which adjusts the sample distribution according to the more reliable data.

The more reliable data used for calculating such factors should, of course, be based upon the very same population represented by the sample. Each of the characteristics chosen to form the categories should be fairly highly correlated with statistics of interest.

I have examined the effect of two alternative sets of post-stratification factors. Both were adjusted to Current Population Survey data. The categories used in each appear in Table 3 and in Table 4. The first set adjusts the data according to the CPS distribution of households by race, residence, size, and income. The second set adjusts the data to the distribution of persons by race, sex, and age. I formed the latter set by trying to group sample persons with similar health characteristics. I also considered the weighted and unweighted number of cases per cell. (Nonresponse adjustment should be performed before post-stratification adjustment. Thus the post-stratification factors used with the data presented in Table 2 differ from those given in Table 3.)

Table 5 presents the effect of the use of post-stratification adjustment on our data. This table does not indicate that there was any great change in the data as a result of using either of the sets of post-stratification adjustment factors. Again however, I do not intend to suggest that these specific results would occur if such adjustments were used with any or all similar data.

In order to definitively determine the effect of alternative post-stratification adjustments, we would need the results of a complete census using the sample survey questionnaire. We have had to examine the effect of post-stratification by comparing adjusted and unadjusted estimates from a single sample. Also, we were able to look at only two different types of estimates — that of mean number of physician visits and of mean total hospital expense per admission. (Our attempt to measure the impact of post-stratifica-

TABLE 3 Original post-stratification adjustment categories and factors

CHARACTERISTIC				POST-STRATIFICATION ADJUSTMENT FACTOR	PERCENT OF SAMPLE HOUSEHOLDS	
Race	Residence	Family Size	Household Income		Unweighted, Unadjusted	Weighted, Adjusted
White	SMSA	1	Under $3000	1.622	6.0%	5.9%
White	SMSA	1	$3000 plus	1.191	5.6	8.0
White	SMSA	2+	Under $3000	1.136	3.3	2.5
White	SMSA	2+	$3000-14999	1.160	18.6	29.8
White	SMSA	2+	$15000 plus	1.181	5.6	12.3
White	NonSMSA	1		0.967	5.9	5.7
White	NonSMSA	2+	Under $3000	1.080	3.3	2.7
White	NonSMSA	2+	$3000-14999	0.750	21.9	18.4
White	NonSMSA	2+	$15000 plus	1.147	2.4	3.8
Nonwhite	SMSA	1	Under $3000	1.200	3.3	1.2
Nonwhite	SMSA	1	$3000 plus	0.714	2.2	1.0
Nonwhite	SMSA	2+	Under $3000	0.714	4.3	1.0
Nonwhite	SMSA	2+	$3000-14999	0.600	13.6	4.1
Nonwhite	SMSA	2+	$15000 plus	1.167	0.8	0.7
Nonwhite	NonSMSA	1		1.000	0.6	0.6
Nonwhite	NonSMSA	2+		0.917	2.6%	2.2%

TABLE 4 Alternative post-stratification adjustment categories and factors

CHARACTERISTIC			POST-STRATIFICATION ADJUSTMENT FACTOR	PERCENT OF SAMPLE HOUSEHOLDS	
Race	Sex	Age		Unweighted, Unadjusted	Weighted, Adjusted
White		0 to 5	1.0344	6.5%	8.7%
White		6 to 11	0.9699	7.9	10.2
White		12 to 17	0.9563	8.3	10.2
White	Male	18 to 29	1.1487	5.0	7.8
White	Female	18 to 29	1.1417	5.5	8.3
White	Male	30 to 44	1.0336	5.1	7.3
White	Female	30 to 44	1.0036	5.4	7.5
White		45 to 54	1.1228	6.7	10.3
White		55 to 64	1.0330	6.6	8.3
White		65 to 74	1.0245	6.5	5.5
White		75 plus	1.1384	4.2	3.5
Nonwhite		0 to 8	0.7906	7.3	2.8
Nonwhite		9 to 17	0.6803	8.4	2.6
Nonwhite	Male	18 to 44	0.9781	3.8	2.0
Nonwhite	Female	18 to 44	0.7465	6.0	2.3
Nonwhite		45 to 64	0.7459	4.7	2.0
Nonwhite		65 plus	0.7410	2.3%	0.8%

tion was further confounded by the fact that noninterview adjustment usually would be performed first; while we have had to consider the effect of each separately. Had we computed estimates using combinations of noninterview and post-stratification adjustment, some combination of the two might have interacted in such a way that such adjustments would have had a bigger effect than either individual adjustment would have suggested.) This information on the effect of post-stratification adjustment, limited though it is, is a useful first step in developing a more thorough investigation into the expected effect of such adjustment on different types of data.

Despite the fact that it is difficult to assess the effect of post-stratification on the data and even more difficult to predict what its use would mean to other surveys, I would suggest that it be done. Post-stratification is an extremely inexpensive procedure and should result in at least some small improvement in the data. The choice of categories depends upon the nature of the survey, since the categories should be delineated by characteristics correlated with important estimates in the study.

TABLE 5 Effect of post-stratification adjustment on distribution of sample persons, on estimates of mean number of physician visits for persons seeing a physician, and on estimates of mean hospital expenditure per admission

	PERCENT OF WEIGHTED SAMPLE PERSONS			MEAN PHYSICIAN VISITS PER PERSON			MEAN EXPENDITURE PER ADMISSION		
	Without Post-Stratification Adjustment	With Post-Stratification Adjustment		Without Post-Stratification Adjustment	With Post-Stratification Adjustment		Without Post-Stratification Adjustment	With Post-Stratification Adjustment	
		Original Categories	Alternative Categories		Original Categories	Alternative Categories		Original Categories	Alternative Categories
Demographic									
Age of oldest family member									
Less than 65 years	86.8%	86.2%	86.5%	5.4	5.4	5.4	$616	$640	$616
65 years or more	13.3	13.8	13.5	7.7	7.8	7.7	830	863	831
Family income									
Nonpoor	75.7	77.0	77.1	5.5	5.5	5.6	649	674	647
Poor	24.3	23.0	22.9	6.3	6.5	6.3	685	715	691
Race									
White	83.8	87.9	87.5	5.7	5.7	5.7	652	684	650
Nonwhite	16.2	12.1	12.5	5.9	6.0	5.9	702	688	730
Residence									
Rural nonfarm	25.8	24.5	26.5	5.4	5.4	5.4	544	557	540
Rural farm	7.8	6.8	7.9	5.5	5.6	5.5	555	575	566
SMSA central city	29.1	29.8	27.7	6.0	6.0	6.0	706	727	711
SMSA other urban	23.3	26.9	23.8	5.5	5.6	5.5	920	910	915
Urban nonSMSA	13.9	12.1	14.2	6.1	6.1	6.2	447	444	447
Perceived and Evaluated Health									
Perception of health									
Excellent	37.0	37.8	37.4	3.5	3.6	3.6	467	488	472
Good	43.3	43.0	43.1	5.2	5.3	5.3	600	609	596
Fair	12.7	12.4	12.7	8.9	9.1	9.0	670	698	682
Poor	3.8	3.9	3.8	14.1	14.2	14.2	816	868	809
Number of diagnoses									
One	27.9	28.2	28.0	3.8	3.8	3.8	605	626	600
Two	17.5	17.6	17.6	5.6	5.6	5.7	573	573	574
Three	8.9	9.2	9.2	8.1	8.0	8.0	543	563	540
Fouror more	8.6	9.0	8.9	11.6	11.6	11.5	836	883	839
Total	100.0%	100.0%	100.0%	5.7	5.7	5.7	$658	$685	$658