NONRESPONSE IN DUAL FRAME SURVEYS BASED ON AREA/LIST AND TELEPHONE FRAMES

Monroe G. Sirken and Robert J. Casady
National Center for Health Statistics

Introduction

When the available sampling frame is complete but costly to use, efficiency considerations suggest the joint use of another less costly incomplete frame. The National Health Interview Survey (NHIS), for example, is currently based on an area/list household frame that covers the entire target population, but it is expensive to use because the data are collected by a personal interview. Efficiency considerations suggest the joint use of a telephone frame that covers only the target population with telephones, but which tends to be less expensive because the data are collected by a telephone interview.

Beginning with the work of Hartley [1962] more than two decades ago, substantial progress has been made in designing dual frame surveys from the viewpoint of sampling efficiency. For a fixed total cost, closed formulas have been derived for minimizing sampling errors by optimizing the sample allocation to the two frames. These formulas take into consideration frame differentials in the data collection costs, sampling errors and coverage completeness. In this paper, the earlier work is extended by adding a fourth factor, namely frame differentials in nonresponse rates.

Casady, Snowden and Sirken [1981] recently explored the possible savings of redesigning the NHIS as a dual frame survey that would be based on an area/list frame, and a frame of randomly dialed telephone numbers. In this connection, the work of Lund [1968] was especially helpful. The optimal sample allocations to the telephone frame in the NHIS ranged from about 65 to 80 percent depending on the particular health variables involved. Assuming a cost structure for the two frames, which as Groves and Lipkowski [1982] have indicated is not yet completely understood, it was demonstrated that the sampling errors would be consistently lower for the dual frame design than for the current NHIS design if the optimum allocation were used.

In contrast to sampling errors, the nonsampling error effects of dual frame designs have received relatively little attention. It would be especially important to take into account nonresponse and other types of nonsampling errors, that are likely to be greater for the less costly frame. For example, current experience indicates that nonresponse rates in telephone interviews are likely to be substantially greater than the personal interview nonresponse rates which currently are about 5 percent in the NHIS. Casady, Snowden and Sirken concluded that it would be premature to proceed with the redesign of a dual frame NHIS until the nonresponse effects of the proposed NHIS dual frame design were investigated.

Consequently, earlier studies, which were limited to the sampling error effects of a dual frame NHIS design were expanded to investigate the dual frame design effects for both sampling errors and nonresponse. Currently, a statistical model is in the process of being developed which will analyze the combined effects of dual frame designs on sampling errors and biases due to nonresponse. The findings in this preliminary report, however, are based on an empirical study of the effect of varying the proportion of the sample allocated to the telephone frame on the nonresponse rates and sampling errors in a dual frame NHIS.

Findings

Both sampling errors and nonresponse rates in dual frame surveys depend on the fraction of the sample allocated to each frame. Assuming different costs and sampling errors for the frames, the allocation that would be optimal for minimizing sampling errors would not be optimal for minimizing nonresponse unless the nonresponse rates were about the same for both surveys. Clearly, if the response rates were unequal, the single frame with the higher response rate would be optimal for minimizing nonresponse rates regardless of the allocation level to the frames that would be optimal for minimizing sampling errors.

The empirical findings in Table 1 refer to three different combinations of telephone and personal interview response rates. The three option sets are based on telephone response rates of 80, 85 and 90 percent, respectively, and a personal interview response rate of 95 percent in all option sets. For each option set, the table shows the effect of varying the percentage of the sample allocated to the telephone frame on sampling errors and on nonresponse in a dual frame NHIS. For a hypothetical variable which is a composite of the variables studied by Casady, Snowden and Sirken, the sampling error effects are expressed as ratios where the numerators are the expected sampling variances for the proposed dual frame NHIS, and the denominators are the existing sampling variances for the current single frame NHIS. The nonresponse effects are expressed as response rates.

The overall findings are essentially the same for all option sets. The sampling error reductions are maximized by allocating about 70 percent of the sample to the telephone frame, and the dual frame design effects become progressively less

Table 1

Effects of Sample Allocations and Telephone Response Rates on Sampling Errors and Response Rates In a Dual Frame NHIS

	Telephone Response Rate					
	80%		85%		90%	
Percent of Sample Allocated to the Telephone Frame	Sampling Error Effects	Response Rate	Sampling Error Effects	Response Rate	Sampling Error Effects	Response Rate
0	1.00	.95	1.00	.95	1.00	.95
10	.97	.93	.97	.94	.96	.94
20	.93	.92	.92	.93	.91	.94
30	.90	.90	.88	.92	.87	.94
40	.87	.89	.86	.91	.84	.93
50	.85	.88	.83	.90	.81	.92
60	.84	.86	.81	.89	.79	.92
70	.83	.84	.80	.88	.78	.92
80	.87	.83	.84	.87	.92	.91
90	1.08	.82	1.05	.85	1.02	.90

advantageous as the allocation deviates in either direction from the optimal. On the other hand, the dual frame response rate is maximized by allocating 100 percent of the sample to the area/list frame, and it declines monitonically with increases in the proportion of the sample allocated to the telephone frame.

Although the proportion of the sample allocated to the telephone frame that is optimal (70 percent) for maximizing the sampling error reductions is invarient across the option sets, the design effects favor the option sets with the higher telephone response rates since the actual sample size is directly proportional to the response rate. For example, at the optimal allocation level, the reductions in sampling error effects are .83, .80, and .78 and the response rates are .84, .88, and .92 for the option sets that are based on telephone response rates of 80, 85, and 90 percent, respectively.

Discussion

In establishing the precision requirements for surveys, nonresponse rates, are readily available as survey by-products and often serve as surrogate measures of nonresponse biases, which are rarely known. Although the relationship between the two measures is tenuous at best, the use of one for the other may be partially justified on the grounds that the larger the nonresponse rate the larger the possible nonresponse biases. Consequently, most survey agencies prescribe minimum response rates that they will tolerate in their surveys. For example, the response rate in the currently designed single frame NHIS is about 95 percent. In considering the possibility of redesigning NHIS as a dual frame survey, it seems unlikely that a response rate under 90 percent would be acceptable.

If $\pi_i > 0$ $(i = 1, 2)$ represents the response rates in the area/list and telephone frames, respectively, and ρ represents the fraction of the dual frame sample that is allocated to the telephone frame, then the response rate in the dual frame survey is given by

$$R = (1 - \rho)\pi_1 + \rho\pi_2 . \qquad (1)$$

It follows that for fixed $\pi_1 = \pi_1^*$ and $\pi_2 = \pi_2^*$ and a specified minimum acceptable response rate, say R^*, the maximal permissible allocation of the dual frame sample to the telephone frame consistent with R^*, π_1^* and π_2^*, is given by

$$\rho^* = \begin{cases} (R^* - \pi_1^*) \;/\; (\pi_2^* - \pi_1^*) & \text{if } R^* > \pi_2^* \\ 1 & \text{if } R^* \le \pi_2^* \end{cases} \qquad (2)$$

when we assume $\pi_1^* > \pi_2^*$. A minimal acceptable response rate $R^* = .90$ in the dual frame NHIS implies $\rho^* = .50$ and $.33$ for the option sets in Table 1 that are based on telephone response rates, $\pi_2^* = .85$ and $.80$, respectively. The sampling error gains are reduced accordingly. At the uppermost allocation levels consistent with $R^* = .90$, the sampling error effects are .83 and .90 for the $\pi_2^* = .85$ and .80, respectively. At the optimal allocation level, $\rho = .70$, however, the sampling error effects would have been .80 and .83, respectively.

On the other hand, $R^* = .90$ does not imply an upper limit $\rho^* < 1$ for the option set based on a telephone response rate $\pi_2 = .90$. Hence, the optimal allocation of sample $\rho = .70$ that maximizes sample error reductions can be realized. At this optimal level, the sampling error effect is .78 and the dual frame response rate is $R = .92 > R^*$. Two conclusions for a dual frame NHIS design are drawn from these empirical findings.

(1) A dual frame design could probably not be justified unless the telephone response rate was at least 80 percent.

(2) It would be cost effective to obtain the highest reasonable telephone response rates since these rates affect both the sampling errors and the response rates in the dual frame NHIS.

References

Casady, R. J., Snowden, C. B., and Sirken, M. G., [1981] "A Study of Dual Frame Estimators for the National Health Interview Survey", Proceedings of the American Statistical Association, Survey Research Methods, pp. 444-447.

Groves, Robert M. and Lipkowski, James M., [1982] "Alternative Multiple Frame, Dual Mode Survey Design", to appear in Proceedings of the American Statistical Association, Survey Research Methods.

Hartley, H. O. [1964], "Multiple Frame Surveys", Proceedings of the American Statistical Association, Social Statistics Section, pp. 203-206.

Lund, R. E. [1968], "Estimators in Multiple Frame Surveys", Proceedings of the American Statistical Association, Social Science Section, pp. 282-268.