

The Subsampling of Nonrespondents on the 2004 General Social Survey

Tom W. Smith

National Opinion Research Center/University of Chicago

April, 2006
June, 2006 Revised

GSS Methodological Report No. 106

1. Background

Since the 1940s, subsampling has been recognized as a method for addressing specific challenges from unit nonresponse. Subsampling (also called a two-phase design or double sampling) involves selecting a portion of nonresponding cases from the original sample at the end of phase 1 and conducting an intensive, and sometimes more tailored, follow-up on the selected cases during phase 2. This approach enables a survey organization to focus its resources on the subsampled cases.

Subsampling was first introduced by Morris Hansen and William Hurwitz (1946). It has been a standard part of the survey sampling repertoire since then (Cochran, 1977; Deming, 1953; Groves, 1989; Groves and Couper, 1998; Hansen, Hurwitz, and Madow, 1953; Kish, 1965; Thompson, 1992). The use of subsampling has increased in recent years, primarily due to the need to address the realities of decreasing response rates. As Groves (2003) notes, “Two-phase designs are increasingly attractive to survey researchers in the U.S. because they offer a way to control the costs at the end of a data collection period.”

Concentrating on a subsample of the more difficult-to-obtain cases allows resources to be focused on this smaller number of cases. This includes additional interviewer attempts, the use of the highest-performing interviewers, and especially the utilization of “converters” who specialize in gaining interviews from initially reluctant respondents. This should result in a higher response rate than if the same total effort had been dissipated across the larger, full sample of nonrespondents. Also, uniformly and fully pursuing all the subsampled nonrespondents may reduce nonresponse bias. In the absence of unlimited time and funds, the interviews successfully obtained from the subsample – cases that may otherwise not have been obtained without an intensive focus of resources -- may reduce nonresponse bias (Elliott, Little, and Lewitsky, 2000; Groves, 2003).

Subsampling does complicate the sample design and reduces the efficiency of the sample. Subsampled responses must be weighted to adjust for the fact that subsampled respondents are “representing” other nonrespondents. Both working and completing

fewer cases and the variability that the weight itself introduces reduces the effective sample size from what it might have been if all of the nonrespondents at the end of phase 1 had been fully pursued rather than subsampled. Moreover, the lower the subsampling rate and, as a result, the larger the weight applied to these cases, the greater will be the reduction in the effective sample size.

The 2004 General Social Survey (GSS)(Davis, Smith, and Marsden, 2005) adopted a two-phase, subsampling of nonrespondents design. This paper 1) describes other recent uses of nonrespondent subsamples, 2) provides an example of this approach to demonstrate its use, 3) illustrates how response rates are calculated when using such a design, 4) presents the outcome from the 2004 GSS, 5) analyzes selected differences by phase, and 6) presents the weights that need to be used on the 2004 GSS.

2. Three Recent Examples of Subsampling to Handle Unit Nonresponse

A number of major data collection efforts have incorporated subsampling as an integral part of their design to maintain the integrity of the data collection while managing costs. Three are cited below.

1. *The American Community Survey and the Census Supplementary Surveys (Census Bureau):* The American Community Survey (ACS) is the largest data collection effort currently underway that involves subsampling. The Census Bureau collects ACS data in continuous, three-month cycles, with a new sample drawn each month. In the first month of a given sample, questionnaires are mailed to the sample households; advance letters, reminder cards and a second mailing are sent to motivate response. In the second month, the Census Bureau follows up with telephone interviews to nonrespondents. Once mail and telephone contacts have failed to elicit response, the Census Bureau selects a subsample of nonrespondents and conducts personal visits to these households during the third month.

The Census 2000 and 2001 Supplementary Surveys were conducted to test the feasibility of proposed ACS methods, including the use of subsampling to enhance response rates. These two years of data collection yielded an average weighted

household-level response rate of 95.9%: mail returns accounted for 51.7% of the sample; telephone interviews accounted for 8.3%. After weighting, personal interviews conducted with the subsampled population represented 36.0% of the sample (Griffin, 2002; Smith, 1998; U.S. Census Bureau, 2001a; 2001b; 2002).

2. *The Chicago Health and Social Life Survey (NORC Population Research Center/University of Chicago)*: In 1995, NORC used a subsample for the Chicago Health and Social Life Survey (CHSLs). In addition to reducing the targeted number of completed interviews, NORC drew a subsample of 465 nonresponding cases at an approximate rate of 1 in 4, dropping three-fourths of the cases and intensifying resources on the subsampled number. Interviewers succeeded in obtaining responses from 40 percent of the subsampled cases (a rate similar to that obtained in the National Survey of Family Growth; see below). The weighted subsampled responses enabled NORC to increase the overall response rate from 64 percent to 71 percent (NORC, 1996).

3. *The National Survey of Family Growth (National Center for Health Statistics/University of Michigan)*: The National Survey of Family Growth (NSFG), conducted by the National Center for Health Statistics, provides national estimates on a range of factors affecting pregnancy and birth rates. Among the challenges identified for Cycle 6 of this survey, which included males and females between the ages of 15-44, was the amount of interviewer effort necessary to gain respondent cooperation and obtain adequate response rates.

In addition to incorporating a two-phase design (11 months for the first phase; 1 month for the second phase), the survey employed a model to determine the “obtainability” of specific cases based on a range of factors, including the type of housing, the age of the respondent, the number of contacts made, respondent reaction at time of contact, etc. These indicators were used to direct interviewers to easier cases in the first phase, and then used again as factors to identify cases for a second phase of subsampled nonrespondent cases. Moreover, some of the strategies used in the first phase for contacting respondents were changed for the second phase on the premise that the approaches used in the first phase were insufficient to motivate participation from the

nonrespondents now selected for the second phase subsample. Changes made in the second phase included, for example, increasing the use of proxy respondents during screening; employing the interviewers who were most productive in the first phase; and making adjustments in the type and amount of incentives.

Using this combination of approaches, NSFG succeeded in obtaining a response rate in the first phase of 64% and a response rate of 40% in the second phase, yielding a combined response rate of 78-79%. Importantly, the approaches used in the second phase tended to yield responses from different population groups than those that were most likely to participate in the first phase (e.g., the first phase attracted more teenagers; the change in approach during the second phase attracted older respondents). The use of subsampling in this instance thus appears to have reduced nonresponse error (Groves, 2003; Groves and Heeringa, 2004).

3. Example

To illustrate how subsampling operates, suppose 6,200 cases are released to the field (Table 1). This example resembles the design utilized for the 2004 GSS. This would be a larger, initial sample than would have been employed without the use of subsampling. This greater number of cases maintains the total target sample size, compensating for the phase 1 nonrespondents that are not followed up after subsampling. Next, suppose that 80.6% percent of the released cases turn out to be eligible, occupied housing units. Suppose, too, that 50 percent of the 5,000 eligible households from phase 1, for 2,500 completed interviews, are interviewed. Then assume also that a subsample of the nonrespondents is selected at a rate of 50 percent. That would mean that:

$$(\frac{1}{2})(5,000 - 2,500) = 1,250$$

cases would be in the nonrespondent subsample. With the extra effort taken for the subsample, assume that about 40 percent or 500 of those subsampled are eventually converted and completed in phase 2. Overall, 2,500 plus 500 = 3,000 cases (the standard GSS target sample size since the biennial, double-sample design was adopted in 1994) would be completed.

Table 1. *Subsampling Scenario for GSS*

Item	Total
Initial sample	6,200
Expected eligibility rate	80.6%
Sample of eligible HUs	5,000
Expected completion rate, phase 1	50%
Expected completed interviews, phase 1	2,500
Expected incompletes, first pass	2,500
Subsampling rate for incompletes	50%
Subsampled cases for phase 2	1,250
Interview completion rate, phase 2	40%
Expected completed interviews, phase 2	500
Total completed interviews	3,000

4. Calculating the Response Rate¹

When using a nonrespondents, subsampling design, weights must be used not only in the substantive analyses, but also in the computation of response and other outcome rates. The response rate is defined as

$$r = \frac{\sum_{\text{completes}} w_i}{\sum_{\text{eligibles}} w_i}.$$

¹ Weighting section adapted from Harter, Wolter, and Scheuren, 2003.

When weights are constant, as for past rounds of GSS, $w_i=w$ and the response rate simplifies to

$$r = \frac{\sum_{\text{completes}} w}{\sum_{\text{eligibles}} w} = \frac{\# \text{completes}}{\# \text{eligible}} = \frac{c}{e}.$$

For the subsampling design, the weights for the subsampled cases are the product of the original phase-one sampling weights and the inverse of the subsampling probability. The response rate becomes

$$r = \frac{w_1 c_1 + k w_1 c_2}{w_1 e_1 + k w_1 e_2} = \frac{c_1 + k c_2}{e_1 + k e_2} = \frac{c_1 + k c_2}{e},$$

where c_1 is the number of completes on the first pass, e_1 is the number of eligibles on the first pass, w_1 is the constant original weight, c_2 is the number of completes in the subsample, e_2 is the number of eligibles in the subsample, and $1/k$ is the subsampling probability.

The new weighted response rate is simple and analogous to response rates computed for prior rounds of GSS. In fact, the scenario in Table 1 above yields a 70% weighted response rate, comparable to the 70% response rate achieved in the 2002 round of GSS.

$$r = \frac{c_1 + k c_2}{e} = \frac{2500 + (2 * 500)}{5000} = .70$$

Furthermore, the weighted response rate is consistent with NORC's statistical standard for computing response rates (Harter and Halverson (2001)), as well as the AAPOR (2006) standards. Notice that the numerator, $c_1+k c_2=3500$, is an estimate of the number of cases that would have been completed if the full original sample had been worked extensively with unlimited resources. Similarly, the denominator, of the weighted rate, $e_1+k e_2=e=5000$, gives the number of eligible cases in the full original sample.

5. Subsampling on the 2004 GSS

For the 2004 GSS at the end phase 1 of the preliminary field period (i.e. after about ten weeks), there were 1440 out-of-scope cases (not housing units, vacant, etc.), 2162 completed cases, 143 partial cases and appointments, 144 final nonrespondents, and 2171 temporary nonrespondents. The temporary nonrespondents were sampled at 50% and 1086 were retained in the study and 1085 were eliminated. The retained subsample cases and the partial/appointment cases were then pursued for approximately another 10 weeks. Ultimately 2812 cases were obtained. The response rate was 70.4% (Davis, Smith, and Marsden, 2005).

6. 2004 GSS Comparisons

Cases completed during the two phases of the 2004 GSS data collection are examined to 1) identify variables on which weighting the data to take subsampling into account would make a substantive difference, 2) point to variables for which the use of subsampling reduces nonresponse bias, and 3) indicate groups for which subsampling estimates are less efficient due to the use of subsampling. Table 2 compares the demographics of cases collected in phase 1 (the initial sample) and phase 2 (the subsample of nonrespondents). There are no statistically significant differences by phase on gender, race, Hispanic ethnicity, marital status, or number of children ever born. The phase 2, nonrespondent cases are more likely to live in large cities and the Northeast, to be Catholic, to be full-time employees and less than 65, to have a high household income and to not report their household income, to have a college degree, and to be Democrats.

7. GSS 2004 Weights

Two weighting factors need to be considered when using the 2004 GSS, the number of adults in the household and the subsampling of nonrespondents.

a. Adults

The full-probability GSS samples used since 1975 are designed to give each household an equal probability of inclusion in the sample. (Call this probability P_h .) Thus for household-level variables, the GSS sample is self-weighting. In those

households which are selected, selection procedures within the household give each eligible individual an equal probability of being interviewed. In a household with n eligible respondents, each has probability P_h of being in a selected household, and $1/n * P_h$ of actually being interviewed. Persons living in large households are less likely to be interviewed, because one and only one interview is completed at each preselected household. For person-level variables, the simplest way to compensate would be to weight each interview proportionally to n , the number of eligible respondents in the household where the interview was conducted. N is the number of persons 18+ (ADULTS) in the household. A discussion of the weight as well and a post-stratification variant of weighting by ADULTS appears in GSS Methodological Report No. 3 (Stephenson, 1978).

b. Subsampling of Nonrespondents

Due to the adoption of the nonrespondent, subsampling design described above, a second weight must be employed when using the 2004 GSS. One possibility is to use the variable PHASE (values 1 for phase 1 cases and 2 for phase 2 cases) and weight by it so that the subsampled cases were properly represented. If one wanted to maintain the original sample size, one would weight by $PHASE * 0.87258$. This weight would only apply to 2004 and would not take into account the number of adults weight discussed above. As such, it would be appropriate for generalizing to households and not to adults.

A second possibility is to use the variable WT2004.2 This variable takes into consideration a) the subsampling of nonrespondents and b) the number of adults in the household. It also essentially maintains the original sample size. In years prior to 2004 a one is assigned to all cases so they are effectively unweighted. To adjust for number of adults in years prior to 2004, a number of adults weight would need to be utilized as described above.

A third possibility is to use the variable WT2004NR. It is similar to WT2004, but adds in an area nonresponse adjustment. Thus, this variable takes into consideration a) the subsampling of nonrespondents, b) the number of adults in the household, and c) differential nonresponse across areas. It also essentially maintains the original sample

² With the release of the 1972-2006 GSS data this weight will be renamed WTSS so it does not have to be renamed with each survey. Likewise, WT2004NR will become WTSSNR.

size. As with WT2004, WT2004NR has a value of one assigned to all pre-2004 cases and as such they are effectively unweighted. Number of adults can be utilized to make this adjustment for years prior to 2004, but no area nonresponse adjustment is possible prior to 2004.

A final possibility, WT7204, adjusts for the subsampling of nonrespondents by using WT2004 for 2004, but replaces the unitary weight in 1972-2002 with a weight adjusting for number of adults. Details on the construction of WT2004, WT2004NR, and WT7204 follow:

WT2004:

W0: Within each NFA, we calculate a probability of selection, n/N . W0 is the reciprocal of this probability of selection (N/n). At this point, each observation stands in for a given number of cases in the NORC sample frame (Davis, Smith, and Marsden, 2005). Because the secondary sample release was only in the urban NFAs, cases in urban NFAs have a slightly higher probability of selection, and thus a slightly lower baseweight, than cases in the urban NFAs.
 $\sum W0 = \text{frame size}$

W1: At the end of Phase I of data collection, we subsampled the nonresponding cases with a sampling fraction $f=.5$. The selected nonresponding cases then get twice their initial weight, and the unselected nonresponding cases get no weight.
 $\sum W1 = \text{frame size}$

W2: Next, we adjust the baseweight for eligibility. Not all cases in the frame are truly eligible for the survey: some addresses in our frame are businesses, do not exist or are unoccupied. We use the eligibility rate of the sampled cases to estimate the eligibility rate for the frame. We calculate the eligibility rate at the NFA level. This adjustment sets the weights of the ineligible cases to missing. Cases whose eligibility could not be determined are given fractional eligibility equal to be eligibility rate for their NFA.
Now the sum of the weights is the estimated number of eligible cases (or occupied housing units) in the frame.
 $\sum W2 = \text{estimated eligible cases in the frame} < \sum W1$

We then rescale W3 so that the sum is the total number of completed interviews. This adjustment helps prevent errors that can arise in SPSS and in some procedures in SAS where the sum of the weights is assumed to be equal to the sample size. The relative weights are unchanged by this adjustment.
 $\sum \text{WEIGHT} = \text{number of completed interviews}$

WT2004NR:

W2NR: We next adjust for nonresponse. Weights for responding cases increase by the reciprocal of the response rate, calculated at the NFA level. The responding cases take on the additional weight of the nonresponding cases. The sum of the weights is the same as the previous step: the estimated number of eligible cases in the frame.

$$\sum W2NR = \sum W2 = \text{estimated eligible cases in the frame}$$

W3: To account for the random selection of an adult respondent, this weight is the household-level weight (W2) multiplied by the number of adults in the household. The sum of the weights in this step is the total number of adults in all eligible households in the frame.

$$\sum W3 = \text{estimated adults in eligible cases in the frame} > \sum W2$$

W3NR: To account for the random selection of an adult respondent, this weight is the nonresponse adjusted household-level weight (W2NR) multiplied by the number of adults in the household. The sum of the weights in this step is the total number of adults in all eligible households in the frame.

$$\sum W3NR = \text{estimated adults in eligible cases in the frame} > \sum W2NR$$

$$\sum W3NR = \sum W3$$

We also rescale W3NR so that the sum is the total number of completed interviews. This adjustment helps prevent errors that can arise in SPSS and in some procedures in SAS where the sum of the weights is assumed to be equal to the sample size. The relative weights are unchanged by this adjustment.

$$\sum \text{WEIGHT} = \text{number of completed interviews}$$

WT7204:

For 2004 this weight is exactly the same as WT2004 (see above). The difference is that for 1972-2002 a number of adults weight has been applied (discussed above). This makes this the weight to use for across survey analysis with individuals as opposed to households being the unit of analysis.

8. Conclusion

The subsampling of nonrespondents is a useful sampling design for dealing with the problem of nonresponse and using scarce resources in a more efficient manner. As utilized on the 2004 GSS, it produced a sample with a similar response rate and comparable in quality to that of recent GSSs before the adoption of the two-phase design.

Table 2

Demographic Comparison of Phase 1 and Phase 2 Cases

	Phase 1	Phase2	Prob.
Gender - % Male	46.6	45.4	.561
Race - % Black	12.2	11.8	.352
Hispanic - % Hispanic	9.5	8.0	.066
Religion - % Catholic	24.1	26.5	.001
Marital - % Married	60.8	59.8	.651
Age - % less than 65	85.7	91.0	.000
Community Type - % Central City	16.1	19.5	.000
Region - % Northeast	15.7	22.8	.000
Education - % College Degree	35.2	42.4	.001
Party Identification - % Democratic	32.7	37.6	.001
Income - % \$90,000+	17.0	22.9	.001
- % Refused/DK	12.2	14.4	
Labor-Force Status - % Full Time	52.0	58.4	.008
Number of Children - % 1+	72.8	71.0	.062

Source: 2004 GSS

References

- American Association for Public Opinion Research (2006). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, Michigan: AAPOR. Available at:
www.aapor.org/default.asp?page=survey_methods/standards_and_best_practices
- Davis, James A.; Smith, Tom W.; and Marsden, Peter V., General Social Survey, 1972-2004: Cumulative Codebook. Chicago: NORC, 2005.
- Cochran, W.G. (1977). Sampling Techniques. 3rd edition. New York: John Wiley & Sons.
- Deming, W.E. (1953). "On a Probability Mechanism to Attain an Economic Balance Between the Resultant Error of Response and the Bias of Nonresponse," *Journal of the American Statistical Association*, **48**, No. 264, pp. 743-772.
- Elliott, M.R.; Little, R.J.A.; and Lewitsky, S., "Subsampling Callbacks to Improve Survey Efficiency," *Journal of the American Statistical Association*, 95 (2000), 730-738.
- Griffin, D. H. (2002). Measuring Survey Nonresponse, by Race and Ethnicity. U.S. Bureau of the Census, Washington, D.C.
- Groves, Robert M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley & Sons.
- Groves, Robert M. et al., (Eds.) (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley.
- Groves, Robert M. and Heeringa, Steven G., "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Nonresponse and Costs," Paper presented to the Conference on Statistical Methods for Attrition and Nonresponse in Social Surveys, London, May, 2004.
- Groves, Robert M., et al (2003). "Using Process Data from Computer-Assisted Face to Face Surveys to Help Make Survey Management Decisions." Paper prepared for presentation at the 2003 Meetings of the American Association for Public Opinion Research.
- Hansen, M. and Hurwitz, W. (1946). "The Problem of Non-Response in Sample Surveys," *Journal of the American Statistical Association*, Vol. 41, No. 236.
- Hansen, M.H., Hurwitz, W.N. and Madow, W. G. (1953). *Sample Survey Methods and Theory, Vol. 1*. New York: John Wiley & Sons.
- Harter, R and Halverson, M. (2001). "NORC Statistical Standard 15 – Calculation of Response Rates."

Harter, Rachel; Wolter, Kirk; and Scheuren, Fritz, "Subsampling Nonrespondents in the 2004 General Social Survey (GSS): Technical Approach," NORC report, August, 2003.

Kish, Leslie. (1965). *Survey Sampling*. New York: John Wiley & Sons.

National Opinion Research Center (NORC) (1996). "Sampling Design for the CHSLS." Internal NORC technical paper.

Smith, Amy Symens. (1998). "The American Community Survey and Intercensal Population Estimates: Where are the Crossroads?" U.S. Census Bureau Population Division Working Paper No. 11. U.S. Bureau of the Census, Washington, D.C.

Stephenson, C. Bruce, "Weighting the General Social Survey for Bias Related to Household Size," GSS Methodological Report No. 3. Chicago: NORC, 1978.

Thompson, Steven K. (1992). *Sampling*. New York: John Wiley & Sons.

U.S. Census Bureau (2001a). "Meeting 21st Century Demographic Data Needs. Report 1: Demonstrating Operational Feasibility." U.S. Department of Commerce, Washington, D.C.

U.S. Census Bureau (2002). "Meeting 21st Century Demographic Data Needs. Report 2: Demonstrating Survey Quality." U.S. Department of Commerce, Washington, D.C.

U.S. Census Bureau (2001b). "Accuracy of the Data." U.S. Bureau of the Census, Washington, D.C.