WEIGHTING THE GENERAL SOCIAL SURVEYS FOR BIAS RELATED TO HOUSEHOLD SIZE

GSS Methodological Report No. 3

C. Bruce Stephenson
Research Assistant

February, 1978

The NORC General Social Surveys have been conducted with two sample designs: the 1972-4 surveys and half of the 1975-6 surveys used a probability sample with quotas at the block level, while the remaining half of the 1975-6 surveys and the 1977 (and subsequent) surveys used a full probability sample in which the respondent was predesignated. Details of the sample designs may be found in Appendix A of the GSS codebooks. Differences between the designs are discussed elsewhere; this report deals with a technique for weighting the surveys to compensate for sampling biases related to size of household. We do not intend to urge users to weight the surveys, but only to lay out the reasons why some may choose to do so, and to briefly describe the necessary procedure and the effect of such weighting. Experience suggests that weighting for household size has only slight effect on analytical results.[1]

The rationale for weighting is quite simple. The full-probability samples used since 1975 are designed to give each household an equal probability of inclusion in the sample. Call this probability P_h. In those households which are selected, selection procedures within the household give each eligible individual 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 simply less likely to be interviewed, because one and only one interview is completed at each preselected household. The simplest way to compensate would be to weight each interview

[1] See, for example, the discussion in Leslie Kish, Survey Sampling (New York: John Wiley & Sons, 1965), p. 400, and in Karen Newman Gaertner, "The Use of AIPO Surveys: To Weight or Not to Weight," in James A. Davis et al, Studies of Social Change since 1948 (Chicago: National Opinion Research Center, 1976), pp. 148-170.

proportionally to n, the number of eligible respondents in the household where the interview was conducted. For all practical purposes, n is the number of adults (persons over 18) in the household.

Two difficulties with this simple scheme motivated the analysis reported here.[2] First, the full-probability sampling procedure is not perfect. Interviews cannot be completed at a substantial proportion (typically over 20 percent) of the preselected households. Many of these failures are due to reasons which are probably unrelated to household size, particularlyrefusal or inability of the selected respondent to participate. Others are due to our inability to contact any household member, even after repeated attempts. One presumes that households which cannot be contacted tend to be small; if so, the known underrepresentation of persons from large households may be offset to an unknown degree by underrepresentation of small households.

The second difficulty is that probability sampling with quotas at the block level was used in the early years of the GSS. The extent to which such samples represent different-sized households is a much more difficult theoretical question. One would expect the prob-with-quotas technique to overrepresent large households, yet underrepresent individuals from large households. Large households are overrepresented because, when an interviewer knocks on the door, they are more likely to contain someone satisfying the quota requirements. At the same time, persons from large households are underrepresented: there may be several persons in the household fitting the quota(s), but only one can be interviewed. Unfortunately, calculation of these expected biases is very

---

[2] See, however, the discussion on page 8.

difficult. If we wish to weight all of the GSS samples to compensate for whatever biases exist in selection from different-sized households, theory cannot help us.

Before proceeding, we establish some terminology. In the following, "household weights" are weights which yield the correct proportion of households with any given number of adults. "Individual weights" yield the correct proportion of persons from households of a given size. There are twice as many adults from two-adult households as there are households containing two adults; and generally, there are n times as many adults from n-adult households as there are households containing n adults. Therefore individual weights for persons from n-adult households should be proportional to n times the household weights for n-adult households. Within each of these categories, weights are standardized when they are adjusted to have a mean of 1.0, so as not to change the sample size.

Note that in general, weighted samples are less efficient than unweighted samples.[3] This means that for precise tests of significance, the weighted number of cases should be <u>less</u> than the number of raw cases. The exact correction factor depends upon the variance of the weights and upon the statistics employed. We will not discuss this correction further, but readers should be aware that weighting incurs a cost in sample efficiency.

In the absence of adequate theory, we shall simply compute weights to get the right answers. We can estimate the true proportions of American households which, in a given year, contained n adults, and we can adjust the GSS proportions to reproduce these. The calculations must be done

---

[3] Kish, p. 400.

separately each year, since average household size has been decreasing rapidly in recent years. There is, however, no reason to believe that biases produced by either sampling technique have changed since 1972. The computed weights should therefore be about the same each year.

The proportions of households with given numbers of adults are not to be found in the published tabulations from the Census or the Current Population Survey (CPS). CPS figures are available for households containing given numbers of <u>persons</u>.[4] To get from number of persons to number of adults, we use a crosstabulation in the GSS of number of persons by number of adults by year. The procedure is best explained by example. In 1976, according to the CPS, .206 of the U.S. households contained one person only. According to the GSS, 100 percent of these included precisely one adult, so we estimate that .206 of the total were one-person, one-adult households. According to the CPS, .306 of all households contained two persons; of these 5.6 percent included one adult and 94.4 percent two adults (GSS). Therefore about .017 of the total households were two-person, one-adult households, and about .289 were two-person, two-adult households. Continuing through household sizes up to seven-or-more persons, we can sum and get estimates of the proportions containing n adults regardless of the number of persons. Estimates are made for each year in the same way. The estimated proportions are given in Table 1.

The assumptions used in this procedure should be made explicit. First, we are assuming that the CPS is not appreciably biased with respect to household size. Second, we are ignoring households which include

---

[4] U.S. Department of Commerce, Bureau of the Census, <u>Currect Population Reports: Population Characteristics: Household and Family Characteristics, March, 1976</u>, Series P-20, np. 311 (August, 1977), Table A, p. 2. Estimates for 1977, not available at the time of this writing, were extrapolated from Table A.

no persons 18 or older. Such households are not a part of the universe represented by the GSS sample; they are presumably included in the CPS tabulations. Their exclusion from the CPS would probably reduce very slightly the proportions of small households. Finally, we are assuming that the GSS yields reasonably accurate estimates of the proportion of m-person households which include n adults. Random error in these estimates is almost certainly the greatest source of inaccuracy in the proportions of Table 1. Fortunately, the proportions which are most affected are those which have small case bases in the GSS, so the least reliable weights will be those which affect only a small number of cases.

We now compute weights for each year, and for each subsample in 1975 and 1976. We have estimated that, for example, .576 of U.S. households in 1976 included precisely two adults. In the full-prob half of the 1976 GSS, .627 of the respondents came from two-adult households, so we assign these respondents a raw household weight of .576/.627 = .919. The raw individual weight is twice this (there are twice as many adults in two-adult households as there are two-adult households), or 1.838. We then attach these raw weights to the GSS cases, compute their respective means, and divide the weights by the means to standardize. Finally, we assign weights of 1.0 to respondents from households containing an unknown number of adults.

Standardized household and individual weights for all samples are presented in Table 2. Note that the household weights for all three full-prob samples remain close to 1.0, particularly for households with small numbers of adults. For all practical purposes, there is no household-size bias in these samples, considered as samples of households. As samples of individuals, they show the expected bias toward persons from small households--a consequence of their being equal-probability for different sized households. The prob-with-quotas weights, as predicted, show these samples to overrepresent large households when considered as sample of households, and to underrrepresent persons from large households when considered as samples of individuals.

The weights of Table 2, applied to the GSS, will produce for each year the distribution of household sizes estimated in Table 1. The trouble with Table 2 is that it simply contains too many numbers. We are trying to correct for biases in the selection of different-sized households, but it would be misplaced effort to insist that every household size be represented in exactly the right proportion each year. Two ways of simplifying suggest themselves.

The first simplification is based on our earlier observation that the weights should remain constant over time. The biases for which we are correcting are certainly a function of sampling technique; but they should not depend upon the changing distribution of household sizes in the population. To collapse the weights of Table 2, we simply (pardon the expression) take a weighted average of the weights for a given sample type and number of adults. In other words, we average over the GSS cases within each sample type and number of adults. These averaged weights, and their correlation with the "exact" weights of Table 2, are reported in Table 3. (Note that the weights of Table 2 are exact only in their correspondence with the proportions of Table 1. The time-averaged weights of Table 3 may be better estimates of the true correction for sampling bias, since they are based on more cases.)

An alternate method of simplifying is possible for the full-prob samples in 1975-7. We have found that the full-prob technique gives very nearly equal probability of selection to different-sized households (as it should, since household size does not enter into the formal selection procedure). We might therefore conclude that the observed small deviations from equal probability are random, and not due to difficulties in contacting households, and that the full-prob samples should be left unweighted when the unit of analysis is the household. This approach implies that the individual weights should be directly proportional to the number of adults in the household. Appropriate individual weights, averaged over the three full-prob samples, are given in Table 4. These weights are in fact the ones mentioned in passing on page 1. We now know that they work well; however, they still apply only to the full-prob samples.

After supplying three alternate sets of weights, we are obligated to give some guidance on whether they are worth the trouble of using, and which are the best to use. The first question is whether weighting makes any difference. We have not made any extensive comparison of weighted and unweighted analytical results in the GSS; we have, however, compared the weighted and unweighted univariate distributions of some important demographic variables. Weights used were the standardized individual weights of Table 3. Naturally, the distribution of household size is affected by these household-size weights. For other demographic variables, we found that weighting the cumulative GSS (1) increased the proportion male by about .01; (2) did not affect the racial composition; (3) increased the proportion Catholic by about .01; (4) did not affect mean occupational prestige; (5) did not affect mean age; (6) increased the proportions currently married and never married by about a percent and a half each;

(7) increased mean family income slightly; and (8) did not affect mean respondent income. In general, then, variables associated with household size are slightly affected by these household-size weights. At least for demographic variables, the univariate effects are small. We suspect that multivariate effects are smaller still.

Some researchers may consider equal-probability representation of different sized households sufficiently important to justify weighting, despite the inefficiency involved and the evidence that weighting makes little difference. In such cases, the first thing to be determined is the unit of analysis. Most survey research focuses on the individual respondent, in which case the "individual weights" given in the tables are appropriate. However, a study of family structure, for example, might require the use of the "household weights." Note that the greater variance of the individual weights means that their use decreases the efficiency of the sample more than does use of the household weights.[5]

The choice between the weights of Tables 2, 3, and 4 depends upon how much the user is willing to assume. In most cases where weighting is desired, the time-averaged weights of Table 3 should serve well. The additional detail in Table 2 is meaningful only if one suspects that the selection biases have been changing, and that these changes are reflected in the numbers of Table 2. The alternate scheme of Table 4, assuming no household-size bias in the full-prob samples, has much to recommend it; however, it provides no guidance for the prob-with-quotas samples.

[5]Kish, pp. 424-433.

TABLE 1

ESTIMATED PROPORTIONS OF U.S. HOUSEHOLDS BY NUMBER OF ADULTS, 1972-1977

Number of Adults	Year					
	1972	1973	1974	1975	1976	1977
1	.207	.211	.221	.234	.251	.265
2	.583	.597	.608	.568	.576	.563
3	.147	.131	.121	.141	.120	.124
4	.047	.049	.037	.039	.042	.036
5	.011	.011	.012	.014	.007	.008
6	.003	.002	.001	.002	.004	.003
7	[a]	[a]	[a]	.001	.001	[a]
8 or more .	.002	[a]	.000	[a]	[a]	[a]

[a]Proportion cannot be estimated: none in GSS sample.

TABLE 2

STANDARDIZED WEIGHTS TO CORRECT FOR HOUSEHOLD-SIZE BIAS IN THE GENERAL SOCIAL SURVEYS, BY SAMPLE[a]

Sample Type	Year	Number of Adults	Household Weight	Individual Weight
Prob-with-quotas . . .	1972	1	1.7120	0.8197
		2	0.9344	0.8947
		3	0.8472	1.2169
		4	0.7932	1.5191
		5	0.7315	1.7511
		6	0.6932	1.9913
		7	[b]	[b]
		8+	0.6312	2.4177
		9 (missing)	1.0	1.0
Prob-with-quotas . . .	1973	1	1.6051	0.7799
		2	0.9272	0.9010
		3	0.8714	1.2702
		4	0.8160	1.5858
		5	0.8340	2.0260
		6	0.8195	2.3889
		7	[b]	[b]
		8+	[b]	[b]
		9 (missing)	1.0	1.0
Prob-with-quotas . . .	1974	1[c]	1.5236	0.7543
		2	0.9367	0.9275
		3	0.8678	1.2889
		4	0.8066	1.5973
		5	0.7162	1.7729
		6	0.7254	2.1546
		7	[b]	[b]
		8+	0.6712	2.6583
		9 (missing)	1.0	1.0
Prob-with-quotas . . .	1975	1	1.8020	0.8814
		2	0.9359	0.9156
		3	0.7679	1.1268
		4	0.7604	1.4877
		5	0.6362	1.5559
		6	0.8828	2.5907
		7	0.4479	1.5337
		8+	[b]	[b]
		9 (missing)	1.0	1.0
Prob-with-quotas . . .	1976	1	1.5375	0.7715
		2	0.9735	0.9769
		3	0.7653	1.1520
		4	0.6116	1.2275
		5	0.6204	1.5565
		6	0.5425	1.6333
		7	0.4393	1.5431
		8+	[b]	[b]
		9 (missing)	1.0	1.0

TABLE 2--Continued

Sample Type	Year	Number of Adults	Household Weight	Individual Weight
Full-prob	1975	1	1.0624	0.5212
		2	0.9671	0.9488
		3	1.0100	1.4864
		4	1.0323	2.0255
		5	1.3176	3.2319
		6	0.8604	2.5324
		7	b	b
		8+	b	b
		9 (missing)	1.0	1.0
Full-prob	1976	1	1.0162	0.5107
		2	0.9201	0.9248
		3	1.2332	1.8592
		4	1.8395	3.6976
		5	1.3726	3.4489
		6	2.6673	8.0424
		7	b	b
		8+	b	b
		9 (missing)	1.0	1.0
Full- prob	1977	1	1.1717	0.5954
		2	0.9570	0.9726
		3	0.9263	1.4121
		4	0.9195	1.8690
		5	0.9432	2.3964
		6	1.0045	3.0627
		7	b	b
		8+	b	b
		9 (missing)	1.0	1.0

[a]See text, pp. 3-4, for definitions of standardized, household, and individual weights.

[b]Not applicable: no GSS cases in this sample have this number of adults.

[c]Two cases in 1974 are coded as having 0 adults; for this analysis they were recoded to 1 adult.

TABLE 3

STANDARDIZED WEIGHTS TO CORRECT FOR HOUSEHOLD-SIZE BIAS IN THE GENERAL SOCIAL SURVEYS, AVERAGED OVER TIME, BY SAMPLE TYPE[a]

Sample Type	Number of Adults	Household Weight	Individual Weight
Prob-with quotas . .	0,1	1.6224	0.7934
	2	0.9379	0.9166
	3	0.8358	1.2238
	4	0.7710	1.5049
	5	0.7203	1.7585
	6	0.6968	2.0402
	7	0.4450	1.5368
	8+	0.6392	2.4658
	9(missing)	1.0	1.0
(Correlation with Table 2 weights:		.98	.97)
Full-prob.	1	1.1049	0.5556
	2	0.9499	0.9545
	3	1.0073	1.5172
	4	1.0985	2.2068
	5	1.1317	2.8322
	6	1.1763	3.5526
	7	b	b
	8+	b	b
	9(missing)	1.0	1.0
(Correlation with Table 2 weights:		.61[c]	.93)

[a]See text, pp. 3-4, for definitions of standardized, household, and individual weights.

[b]Not applicable: no GSS cases in these samples contain this number of adults.

[c]This low correlation is relatively unimportant, since there is so little variation in the household weights for the full-prob samples, whether in Table 2 or Table 3.

TABLE 4

STANDARDIZED INDIVIDUAL WEIGHTS FOR FULL-PROB CASES, ASSUMING EQUAL-PROBABILITY REPRESENTATION OF DIFFERENT-SIZED HOUSEHOLDS[a]

Number of Adults	Individual Weight
1	0.4996
2	0.9992
3	1.4988
4	1.9983
5	2.4979
6	2.9975
9 (missing)	1.0

[a]See text, page 8, for discussion. These weights are simply the number of adults divided by the mean number of adults, which is 2.0016.

NORC
UNIVERSITY OF CHICAGO
NATIONAL OPINION RESEARCH CENTER • 6030 SOUTH ELLIS • CHICAGO, IL 60637 • 312/753-1300

DATE February 17, 1978

MEMORANDUM

TO Friends of GSS

FROM Art Stinchcombe

SUBJECT A Non-obvious Consequence of Stephenson's Study of Weighting

As Kish points out, the more heterogeneous the weights that need to be attached to the cases, the larger the sample you need. He gives a formula on p. 430 for computing approximately the loss, as compared with an equal probability sample, for different weights needed due to unequal probabilities of cases falling in the sample.

Using the weights that Bruce Stephenson gives in Table 4, together with the distribution of households by number of adults in them obtained in Table 1, I have calculated that a sample has to be about 18 percent bigger the way we draw probability samples and then interview one person, than it would be if we could draw samples with a probability of a household falling in our sample proportional to the number of adults in it. That is, our probability samples require us to get samples about 18 percent bigger because they require us to weight by number of eligible adults, when estimating a mean or a proportion.

Because the weights are not as variable with block quota samples, the loss is smaller there. I calculate that using the weights Bruce has derived for block quota samples in the top part of Table 3, a block quota sample has to be about 3 percent larger than one which gives people from different households equal probability of falling in the sample, regardless of the size of their household.

Combining these two results, it turns out then that in order to get the same level of accuracy from a full-prob sample, one has to have about 14-1/2 percent more cases than one has to have if one draws a block quota sample.

Of course, there are other advantages of a full-prob sample which may well more than compensate for its slightly smaller efficiency. Aside from the reasons we all know (e.g., that we do not know the biases in block quota samples from taking easily available people), it should also be pointed out that <u>unless</u> a variable is related to household size, <u>weighting is not necessary</u>, and the relative sizes of the sample necessary calculated above do not apply to the unweighted estimates.

ALS:tt

To All and Sundry   Art Stinchcombe

Improvements of Scales by adding items, where the items are of different efficiencies individually.

The following two columns give the estimated validitiea of scales composed of different numbers of items, if the items are of different original measuring qualities. On the left is a column when the items are"bad", with each item having an estimated validity of .40, so that the expected iteritem correlation or test-retest reliability would be .16. On the right is a column estimating what happens when the items are "good" measurements, with validites of .55, leading to an interitem correlation on the average of .3, or a reliability of the same amount. Roughly speaking, if you increase the measurement efficiency of a variable by a factor of 1.414 (the square root of two), you decrease the sample size necessary to find a given relationship by a factor of 2. So the following tables should be useful in calculating the approximate tradeoff between asking more questions and asking more people those questions.

	ESTIMATED VALIDITIES	
Number of Questions	Interitem Correlations Averaging .16	Interitem Correlations Averaging .3
1	.4	.55
2	.53	.68
3	.60 ← 1/2	.75
4	.66	.79 ← 1/2
5	.70	.83
6	.73	.85
7	.76	.87
8	.78	.88
9	.79	.89
10	.81 ← 1/4	.90
20	.89	.95
50	.95	.98
100	.97	.99

The arrows tell how many items you have to have to halve the sample size you need, or to cut it in a quarter, as compared with the sample size you need to find the relationship between this seale and some other variable when you have only one item.