IMPROVING ON PROBABILITY WEIGHTING FOR HOUSEHOLD SIZE

ANDREW GELMAN
THOMAS C. LITTLE

Introduction

In survey sampling, inverse-probability weights are used to correct for unequal selection probabilities, and poststratification weights are used to correct for known or expected discrepancies between the sample and the population (see, e.g., Kish 1992). In this research note, we consider the effects of these adjustments for household size in telephone polling.

In a survey in which households are sampled at random, and then a single individual is sampled from each sampled household, individuals in larger households have a smaller probability of being selected. The probability of an individual being included in the survey is inversely proportional to the size of the household in a simple random sample of households if individuals within a household are selected with equal probability and there is no nonresponse. However, composition of the sample is also affected by nonresponse. One source of nonresponse is nonavailability—no one answers the phone, or no one receives the message on the answering machine. It seems reasonable to suppose that in a larger household it is more likely that someone will be home to receive the phone call. Another source of nonresponse is refusal to participate in the survey.

Method

To study empirically how nonresponse rates vary by household size, we compare responses from national polls to U.S. Census figures on household size (from the 1990 Public Use Micro Survey data). We analyze the telephone polls conducted by CBS News and the *New York Times* in the months preceding the 1988 U.S. presidential election. (For brevity, we refer to these as CBS polls.) These surveys are of particular interest because, unlike many national polling organizations, CBS uses weights proportional to household size as part of its survey adjustments. We break

ANDREW GELMAN is associate professor in the Department of Statistics, Columbia University. THOMAS C. LITTLE is an associate at Morgan Stanley Dean Witter, New York.

Public Opinion Quarterly Volume 62:398–404 0033-362X/98/6203-0005$02.50

the CBS surveys into two groups: early (three polls conducted more than 80 days before the election, with a total of 4,248 respondents) and late (seven polls conducted in the 2 weeks before the election, with a total of 9,818 respondents). Each of the early polls was conducted during a period of 3–4 days, and each of the late polls was conducted over 2–3 days. We also examine the National Election Study (NES), a survey with in-person interviews of 2,040 respondents, which we would expect to look more similar to the population of U.S. adults.

Results

Comparisons of the surveys to the census appear in table 1. The first two columns of the table show the distribution of number of adults in household from the census, counting by household and by adult, respectively. The remaining columns show the proportion of survey respondents in each category of household size, along with the weighted proportions (computed by multiplying unweighted proportions by number of adults in households, then renormalizing so the total is 1). For the CBS surveys, we also present the weighted averages using the complete CBS weights, which are computed based on number of adults in household, number of telephone lines in household, region of the country, race $\times$ sex, and age $\times$ education, in that order (see Voss, Gelman, and King [1995] for details).

Compared to the census results by household, the CBS surveys include too few households with one adult (e.g., 25.3 percent of respondents in the late CBS polls compared to 34.9 percent of census households) and too many households with three or more adults (e.g., 16.4 percent of respondents in the late CBS polls compared to 9.9 percent of census households). As a result, the weighted results overrepresent adults who live in large households. The results for early and late polls are nearly identical. In contrast, the NES survey overrepresents the large households only slightly, and the weighted results are very close to the census proportions for individuals.

For example, the census tells us that 19.6 percent of adults in the United States live in households with no other adults. For the early CBS polls, the proportion who live in such households is estimated as 24.0 percent from the unweighted data, 11.8 percent when weighting by number of adults, and 12.7 percent using the complete CBS weighting. The late CBS polls give similar estimates (25.3 percent [unweighted], 12.7 percent or 13.4 percent [weighted]), but the NES poll gives estimates of 33.6 percent (unweighted) and 18.3 percent (weighted).

A possible cause of the overrepresenting of large households in the weighted CBS polls is that large households are more likely to have addi-

Table 1. Distribution of Number of Adults in Household by Household and by Person (from the 1990 census) and among Respondents of CBS Telephone Surveys and the NES In-Person Survey Preceding the 1988 U.S. Presidential Election

Number of Adults in Household	Proportion of Households (census)	Proportion of Adults in Each Type of Household (census)	Proportion of Respondents (early CBS polls)			Proportion of Respondents (late CBS polls)			Proportion of Respondents (NES poll)	
			No Weights	Weights (number of adults)	Weights (CBS)	No Weights	Weights (number of adults)	Weights (CBS)	No Weights	Weights (number of adults)
1	.349	.196	.240	.118	.127	.253	.127	.134	.336	.183
2	.552	.622	.576	.567	.558	.582	.586	.578	.532	.579
3	.078	.132	.120	.178	.179	.111	.168	.169	.099	.161
4	.017	.038	.048	.095	.093	.038	.076	.076	.026	.058
5+	.004	.012	.015	.042	.043	.015	.042	.043	.006	.020

NOTE.—Weighted proportions (number of adults) are computed by multiplying unweighted proportions by number of adults in households, then renormalizing so the total is 1. Weighted (CBS) proportions use the CBS weighting scheme, which includes the number of adults weights and also adjustments for the number of telephone lines, geography, and demographics. If sampling all went as planned, the ''unweighted'' proportions from the surveys should match the proportion of housholds from the census, and the ''weighted'' proportions from the surveys should match the proportion of adults from the census. The CBS polls clearly oversample the larger households. The NES respondents match the population of households, and so the NES weighted respondents match the population of adults much more closely.

tional phone lines (and thus be more likely to be included in a random telephone sample), but we found this effect to be minor. The weighted (CBS) columns of table 1, which include weighting for phone lines along with other adjustments, differ only slightly from the weighted (number of adults) columns for the CBS polls, which do not include weights for phone lines.

Recommendation

It is well known that in sampling one individual per household a survey organization will oversample individuals from small households. Weighting by number of adults in household corrects for this consequence of the sampling design, but it does not correct for the opposite effect that large households are easier to reach and will be overrepresented in the sample.

We have found that probability weighting for household size can be effective (for the NES) or worse than unweighted responses (for the CBS polls). So what should a survey analyst do? We recommend an alternative strategy of poststratification on the census totals for the proportion of adults in households with 1, 2, 3, 4, 5+ adults (pooling the last two or three categories for small surveys). For each category, the poststratification weight is computed as the proportion of adults from the census divided by the proportion of survey respondents in that category. For example, for the late CBS polls, the weights for respondents in households with 1, 2, 3, and 4+ households would be 0.196/0.253, 0.622/0.582, 0.132/0.111, and (0.038+0.012)/(0.038+0.015), respectively. Table 2 displays the poststratification weights for the CBS and NES surveys, with the weights renormalized to equal 1 for respondents in households with one adult. By comparison, the table also gives the theoretical weights that would be obtained under a large simple random sample of households.

If weighting or poststratification is performed on other variables, then number of adults in household can be added as an additional variable in the weighting procedure. For example, for the CBS polls, we begin with the CBS weights, then use iterative proportional fitting to match to population totals for region of the country, the demographic variables, and number of adults in household (using the categories 1, 2, 3, 4+).

There are two advantages of performing poststratification in addition to weighting proportional to the number of adults in household. First, and most important, poststratification automatically causes the survey to match the census (if the most recent census is several years old, data from a more recent Current Population Survey can be used instead), whereas weighting by number of adults can seriously overrepresent large households.

Table 2. Poststratification Weights for Late CBS Polls, Early CBS Polls, and NES, Normalized So That the Weight is 1 for Respondents from Households with One Adult

Number of Adults in Household	Poststratification Weights			
	Theory	Early CBS	Late CBS	NES
1	1	1.00	1.00	1.00
2	2	1.32	1.38	2.00
3	3	1.35	1.53	2.30
4+	4.25	0.95	1.20	2.55

NOTE.—If sampling all went as planned, the weights would equal the theoretical values. (The last weight is not exactly 4 because the last poststratification category includes all households with 4 or more adults.) The weights for the higher categories are lower than the theoretical values because the surveys oversampled the larger households.

Second, the poststratification weights (see table 2) are, in fact, less variable than the weights 1, 2, 3, 4, 5, and so on, obtained from household size, which will reduce the standard errors of weighted sample means. This pattern also holds after adjusting for other variables (as we can see by computing for each survey, the coefficient of variation of weights used by CBS, which include weights proportional to number of adults in household, and the coefficient of variation of the weights obtained after poststratification by household size in addition to the CBS adjustments). Poststratification reduces the coefficient of variation of the weights in any given survey from about 63 percent to about 48 percent.

Practical Implications

The weighted CBS polls do not match the census on the distribution of household size, but does this cause problems in practice? We investigate this question by examining the influence of the weighting method on the question of primary interest in the survey—preferences in the presidential election. For each of the CBS surveys, we compute the average response to the presidential preference question, considering four different weighting schemes: (1) no weights, (2) weights proportional to number of adults in household, (3) the CBS weights (which include weights proportional to household size along with other adjustments), and (4) iterative proportional fitting applied to the CBS weights so as to match the census on household size and also to agree with the CBS poststratification vari-

Table 3. Effect of Different Weighting Schemes on the Estimated Support for the Presidential Candidates, Based on the Average Estimated from the Late CBS Polls Conducted during the 2 Weeks Preceding the Presidential Election

	Early CBS Polls				Late CBS Polls			
Response	No Weights	Weights (number of adults)	Weights (CBS)	Poststratified	No Weights	Weights (number of adults)	Weights (CBS)	Poststratified
George Bush	.452	.456	.443	.447	.476	.485	.460	.461
Michael Dukakis	.430	.428	.432	.426	.373	.370	.380	.373
Neither/no response	.118	.116	.125	.127	.151	.145	.160	.165

NOTE.—The following weighting schemes were considered: (1) no weights, (2) weights proportional to number of adults in household, (3) the CBS weights (which may include weights proportional to houshold size along with other adjustments), and (4) iterative proportional fitting applied to the CBS weights so as to match the census on household size and also to agree with the CBS poststratification variables. The table shows that the CBS weights have a noticeable effect, but the treatment of household size is essentially irrelevant for this particular outcome.

ables. Table 3 displays the averages for the early and late CBS polls. The averages for the two sets of polls differ (there was a shift in preference from Dukakis toward Bush during the campaign), but in both cases, the household weights have virtually no effect. In contrast, CBS's geographic/demographic adjustments (included in both the CBS and poststratified weightings) have noticeable effects (see Voss, Gelman, and King [1995] for more evidence of this).

The reason why the household adjustments are inconsequential for this variable is that adults in one-adult households and three-adult households tend to support the Democrats, whereas adults in two-adult households tend to support the Republicans. The main effect of different weighting schemes is to reallocate the weights between the one-adult and three-adult households, and so the effect on the average support for the different presidential candidates is minor in this example (although not necessarily in general). This may be one reason that many major national political polls in the United States do not adjust for household size (see Voss, Gelman, and King 1995).

We emphasize that this work is not meant in any way as a criticism of the CBS polling practices; on the contrary, we are grateful that CBS and the *New York Times* have gathered the information on the number of adults in households that has allowed us to perform this research. We conclude that using weights proportional to the number of adults in the household leads to predictable biases due to nonavailability/nonresponse that can be corrected using poststratification, yielding final weights that are less variable and that more accurately fit the target population.

References

Kish, L. 1992. ''Weighting for Unequal P_i.'' *Journal of Official Statistics* 8:183–200.

Voss, D. S., A. Gelman, and G. King. 1995. ''Pre-election Survey Methodology: Details from Nine Polling Organizations, 1988 and 1992.'' *Public Opinion Quarterly* 59:98–132.