ESTIMATION OF CORRELATED RESPONSE VARIANCE WITHOUT RANDOMIZATION OF INTERVIEWER ASSIGNMENTS

Sandra H. Berry, The Rand Corporation

In their recent report on methods of assessing survey practice Bailar and Lanphier (1978) urge that more attention be paid to the contribution of interviewers to total survey error. The current method for doing this, calculation of the portion of the total variance of each survey response attributable to interviewers, is simple in theory, but requires random assignment of interviews to interviewers. This increases the cost and complexity of fieldwork and may preclude most effective use of the interviewers' talents, thus lowering response rates. Use of a well-known statistical procedure, however, can permit calculation of interviewer variance when assignments are not random, but have known biases. While not as good as calculations based on random assignment, the results can be useful in assessing survey quality.

The method is based on Kish (1962) who estimated interviewer variance using a random effects model for a one-way analysis of variance. This assumes that the interviewers used in the study are a random sample of all possible interviewers. Each interviewer is considered to be a "treatment" given to all members of the group of respondents that are part of the interviewer's workload. Since each interviewer's workload is also assumed to be a random sample of cases, this fits the random effects model and an inference can be made about interviewer effects in general from the set of interviewers who actually worked on the survey.

The total variance of each survey response Y_{ij} over the population of all potential observations is

$$\sigma^2 Y = E(Y_{ij} - \mu)^2 = \sigma^2 A + \sigma^2 \varepsilon$$

so that variance consists of two components: $\sigma^2 A$ is the variance due to interviewers and $\sigma^2 \varepsilon$ is the error variance or the variance due to sampling and simple response variance. The interviewer variance reflects the variance due to each interviewer's misinterpretations, carelessness, etc. The sampling variance reflects the fact that the survey was conducted on a sample of respondents (versus a census). The simple response variance reflects the trial-to-trial variability in any one respondent's answers, uncorrelated with interviewers. Other components of error, such as coding error, also appear in $\sigma^2 \varepsilon$. The unbiased estimate of $\sigma^2 A$ may be found by subtracting the mean square within interviewers from the mean square between interviewers and dividing the result by the number of cases in each interviewer's assignment as shown in Table 1 where

$$n = \sum^{h} \bar{n}, \qquad y = \sum^{h} y_j, \qquad y_h = \sum_{j}^{\bar{n}h} y_h$$

with h = the number of interviewers, n = the sample size, and k = the number of cases in each interviewer's assignment. The value k, rather than simply $\bar{n} = n/h$, is used to correct for variation in the size of interviewer assignments. Thus,

$$k = \frac{n}{h}\left[1 - \frac{\mathrm{var}(n/h)}{(n/h)^2}\right]$$

The between interviewer variance is the average of squared differences between interviewer means for a survey response and the sample mean for that response. Within interviewer variance, the sum of sampling, simple response, and the other error variances is the differences between responses for individual cases in an interviewer's assignment and the mean of those responses in the interviewer's assignment. The interviewer variance is the difference between total variance and within interviewer variance.

The ratio of the estimate of $\sigma^2 A$ to the estimate of $\sigma^2 \varepsilon$ is the estimated proportion of the variance accounted for by interviewers. The MS between divided by the MS within is the F-statistic, used to test significance. However, this is sensitive to the number of interviewers and assignment sizes and is not appropriate for comparisons across studies. The appropriate statistic for such comparisons is the population intraclass correlation coefficient.

$$\sigma_I = \frac{\sigma^2 A}{\sigma^2 A + \sigma^2 \varepsilon}$$

which expresses the proportion of total variance accounted for by the interviewers. This is estimated by

$$\text{est } \sigma_I = \frac{\text{est } \sigma^2 A}{\text{est } \sigma^2 A + \text{est } \sigma^2 \varepsilon}$$

$$= \frac{\text{MS within} - \text{MS between}/k}{\frac{\text{MS within} - \text{MS between}}{k} + \text{MS within}}$$

When cases are assigned randomly to interviewers (as in the above procedure) it is assumed

Table 1

CALCULATION OF UNBIASED ESTIMATE OF $\sigma^2 A$

Source of Variance	Degree of Freedom	Sum of Squares (SS)	Mean Square	Expected Values
Between interviewers	h - 1	$\sum^{h} \frac{y_h^2}{\bar{n}} - \frac{y^2}{n}$	$\frac{SS(\text{between})}{h-1}$	$k(\sigma^2 A) + \sigma^2 \varepsilon$
Within interviewers	n - h	$\sum^{h}\sum^{\bar{n}} y_{hj}^2 - \sum^{h} y_h^2/\bar{n}$	$\frac{SS(\text{within})}{n-h}$	$\sigma^2 \varepsilon$

that background variation in assignments is evenly distributed across interviewers. There are, in fact, problems with this assumption. It is true for all possible sets of random assignments, but for any one selection of assignments (as in a particular survey) the distribution for some assignment characteristics across interviewers could be uneven. However, when assignments are known to be nonrandom, the assumption that respondent characteristics are balanced across assignments is clearly untenable.

This was the case in the 1976 baseline survey for the Los Angeles Electricity Rate Study (ERS). In this personal interview survey of about 2700 households certain assignment criteria were used which were intended to control distribution of cases to interviewers. Assignments were based on the geographic location of the case, on expected difficulty of the assignment, and, to some degree, on race of respondent. Throughout the survey an effort was made to minimize travel between interviewers' homes and their assignments. Further, since many interviews also included the task of enrolling the respondent in an experimental electricity rate plan, more successful interviewers were given more difficult assignments. Finally, reassignment of cases after a first refusal was often based on matching race of interviewers and respondents.

While these assignment policies were in effect during the survey other factors in the field procedures (e.g. fielding cases in one geographic area at a time and the need to balance size of assignments) resulted in most cells of assignments being covered by many interviewers and most interviewers working in many cells. This provided a basis of comparison of interviewers working in similar conditions, although the assignments were not randomly distributed. Because the sources of bias in assignments were known, the method of controlling statistically for the factors that affected assignment was used to improve the quality of variance estimates. This was done by using an analysis of covariance with assignment factors as fixed factors and interviewer influence treated as a random variable. This is a standard way of removing the effect of nuisance factors, thus improving the estimate of the effect of relevant factors.

The calculation of interviewer effects was of particular interest in the ERS baseline survey. Interviewing was carried out by employees of the Los Angeles Department of Water and Power (sponsor of the study) who were mainly drawn from the line maintenance and meter setting divisions. The fact that they were predominantly male and were unused to clerical tasks made them an unusual field staff (see Sudman, 1974). In addition, inexperienced interviewers have a known negative effect on data quality (Rustemeyer, 1977). Finally, the unusual organizational arrangements involved in the use of this staff raised problems of survey control. Concern with assessing the overall quality of interviewer performance led to a desire to estimate interviewer variance as one of a number of quality measures for the ERS survey and to compare it with available estimates of interviewer variance in other surveys.

Interviewer variance estimates were produced for a set of factual variables for which 1970 Census estimates of interviewer variance were available. To be comparable with U.S. Census estimates, continuous variables were recoded into categories and expressed as proportions. For example, interviewer variance for the variable "number of rooms" was calculated once for the continuous variable and for each of eight categorical variables (percent with two rooms, three rooms, etc.).

Multiple regression was used to produce the actual estimates. The interviewer variance estimate for each variable required two equations. Each dependent variable (such as "number of rooms") was first regressed on the dummy variables representing geographic zones, difficulty level of assignment (amount of electricity used category), and race of respondent. A second model was then estimated for the same dependent variable with the assignment variables (same as in first equation) entered as the first step and dummy variables for each interviewer (except one, of course) entered as a second step. This yielded two calculations of regression sums of squares (SS); the SS with only assignment variables was subtracted from the SS with assignment and interviewer variables yielding the between interviewer assignment SS. This is shown in Table 2.

Before presenting the result it should be noted that controlling for non-random assignment factors in this way could lead to an underestimate or an overestimate of interviewer variance.

Table 2

CALCULATION OF INTERVIEWER VARIANCE ESTIMATE

Source of Variation	Degree of Freedom	Sum of Squares (SS)
Assignment variables[a]	c	Regression SS without interviewers
Between interviewer	h - 1	Regression SS with interviewers minus Regression SS without interviewers
Within interviewer	n - (h + c)	Error SS with interviewers (same as Total SS minus Regression SS with interviewers)
Total	n - 1	

[a]Includes dummy variables for geographic location, electricity use category, and race of respondent (c = the number of dummy variables).

If, for example, some unknown factor influenced the assignment of cases and was not controlled in this model, the estimate of variance attributable to interviewers could be increased. If, on the other hand, the interviewer effect was correlated with one of the assignment factors included in the model the estimate of interviewer variance would be decreased.

The estimates produced for the ERS survey are compared with estimates for the 1970 U.S. Census (Tepping and Bailar, 1973) in Table 3. The Census estimation method, which attempts at a conceptual level to estimate the same component of error variance, is so closely linked to clustering of assignments and random assignment of cases that it could not be used for the ERS study. Nevertheless, it provides the best comparison available because variances are estimated for similar items. The ERS variance estimates were converted to relvariances for comparison with the Census. This was done by dividing the interviewer variance by the squared proportion of responses which fell into the category (Bailar, 1976). For comparison with earlier surveys, some

Table 3

INTERVIEWER VARIANCE IN ELECTRICITY RATE STUDY (ERS) BASELINE SURVEY COMPARED WITH 1970 CENSUS AND OTHER STUDIES

Variable Name	ERS Percent of Households[a]	1970 Census Percent of Households	ERS Interviewer Relvariance[b]	1970 Census Interviewer Relvariance	F-Statistic for ERS Variance	Approximate Comparative F-Statistic for Other Studies
Number of rooms			0.01860	---	1.40*	1.46[e]/2.80*[f]
1 Room	0.4	1.4	Not computed	0.62714	Not computed	
2 Rooms	2.8	2.7	0.61538	0.08184	2.03**	
3 Rooms	13.1	10.9	0.07809	0.00000	1.82*	
4 Rooms	14.7	19.0	0.00926	0.00586	1.09	
5 Rooms	19.6	24.7	0.03748	0.00292	1.52*	
6 Rooms	23.1	21.5	0.01031	0.00000	1.18	
7 Rooms	12.7	10.8	0.00496	0.00049	1.05	
8 Rooms	6.9	5.7	0.01261	0.00617	1.06	
9+ Rooms	6.7	3.2	0.03118	0.09479	1.14	
Not Reported	2.0	1.5	3.13636	0.39478	1.37	
Persons in HH			0.00799	---	1.20	0.88/1.08[e]
1 Person	18.8	15.4	0.23851	0.00843	1.21	
2 Persons	30.0	29.3	0.00000	0.00179	0.95	
3-4 Persons	33.7	34.8	0.00000	0.00332	0.96	
5-6 Persons	14.3	16.0	0.00880	0.00087	1.31	
7+ Persons	2.8	4.6	0.00000	0.00658	0.71	
Wage and salary income			568,972.00000	---	1.09	
$2,999 or less	18.1		0.29556	0.01218[c]	5.64**	1.30[g]
$3,000-4,999	5.3		0.02151	0.00000	1.08	
$5,000-6,999	3.8	Not	0.02062	0.00711	1.06	
$7,000-9,999	6.6	Available	0.00000	0.00867	0.99	
$10,000 or more	52.8		0.02164	0.00056	3.01**	
Not reported[d]	16.2		0.48835	0.08275	6.71**	2.21[g]
Tenure						
Owned	65.8	63.5	0.00002	0.00001	1.48*	1.82*[f]
Co-op or Condominium	2.4	1.4	0.00000	0.25748	0.93	
Renter	33.8	33.8	0.00357	0.00035	1.62*	
Not reported	0.1	1.4	0.40000	0.24887	1.05	
Renter pays gas	94.5	Not Available	0.00171	0.00352	1.56*	

*indicates significance at the $\alpha = 0.05$ level.

**indicates significance at the $\alpha = 0.005$ level.

[a]Percent computed as $\frac{\text{\#HH's with that characteristic}}{\text{\#HH's who were asked the question}}$.

[b]Variance for continuous version of the variable is simple variance. Variance for proportions is expressed as a relvariance (see text for definitions).

[c]Census figure is wage and salary income for males only.

[d]Not reported as continuous variable in final complete cases.

[e]Gray (1956), British government social survey on noise in residences.

[f]Kemsley (1960), British government social survey on household expenses.

[g]Hanson and Marks (1958), U.S. Census conducted in 1950.

of which only reported F-statistics, the F-statistics for the ERS survey are also reported. As noted above, these should be interpreted with caution.

In general, ERS interviewer variances are somewhat larger than those for the 1970 Census, as might be expected, since the 1970 Census was a mailed questionnaire with enumerators used only in problem cases. As is usual, the relvariances for both surveys are quite small and the ERS relvariances for each variable are, with few exceptions, within the same range as the U.S. Census. They also follow the same general pattern as the Census, with higher relvariances at the extremes of distributions and for the "not reported" category. When compared with F-statistics from earlier surveys which used personal interviewers, the ERS F-statistics are still within range. For the wage and salary income variable, however, the general statement is not true; the ERS income variable, showed larger interviewer effects than other items in the ERS survey and income items in the Census or the earlier surveys.

The findings from the interviewer variance estimates in the ERS are consistent with other measures of interviewer quality for the study (Berry, 1979). Overall, the ERS interviewers compared favorably with regular interviewing staffs from university survey centers. They achieved high response and low refusal rates, showed little evidence of dishonesty or carelessness, produced few edit problems, and attained low levels of item non-response. The exception to these positive findings was for the income variable. The ERS interviewers were uncomfortable asking for an exact dollar amount for income. This presented the only area of difficulty in interviewer training and continued to be a problem throughout the study. When data quality was examined after the survey this problem appeared in higher levels of item nonresponse compared with other items in the ERS survey and income items in other surveys and higher levels of interviewer variance.

The explanation that interviewers' prior attitudes about the difficulty of collecting income data affected their performance receives support from Sudman et al (1977). They found that interviewer expectations of difficulty had weak effects on levels of reporting and suggested that interviewers who expect a study to be difficult not be hired for the study. It is not surprising that these effects would be more pronounced among inexperienced interviewers (as in the DWP survey) than among Sudman's highly experienced staff.

To summarize, use of the covariance method of calculating interviewer variance allowed estimates to be generated in a situation where they would have been otherwise unobtainable. The estimates, when compared with those for other surveys and combined with other measures of interviewer quality, helped to form a more accurate total assessment of interviewer quality. The concerns raised by the unusual circumstances of the ERS survey made this assessment especially useful.

References

Bailar, Barbara A., "Some Sources of Error and Their Effect on Census Statistics," *Demography*, Vol. 13, 1976, pp. 273-286.

Bailar, Barbara A., and C. Michael Lanphier, *Development of Survey Methods to Assess Survey Practices*, American Statistical Association, Washington, D.C., 1978.

Berry, Sandra H., *Conducting a Survey Using the Client's Staff: Evaluation of Interviewer Performance in the Electricity Rate Study*, The Rand Corporation, R-2223-DWP, September 1979.

Gray, Percy G., "Examples of Interviewer Variability Taken from Two Sample Surveys," *Applied Statistics*, Vol. 5, 1956, pp. 73-85.

Hanson, Robert H. and Eli S. Marks, "Influence of the Interviewer on Accuracy of Survey Results," *Journal of the American Statistical Association*, Vol. 53, 1958, pp. 635-655.

Kemsley, W.F.F., "Interviewer Variability in a Budget Study," *Applied Statistics*, Vol.9, 1960, pp. 122-128.

Kish, Leslie, "Studies of Interviewer Variance for Attitudinal Variables," *Journal of the American Statistical Association*, Vol. 57, 1962, pp. 92-115.

Sudman, Seymour and Norman M. Bradburn, *Response Effects in Surveys*, Aldine, Chicago, 1974.

Sudman, Seymour, Norman Bradburn, Ed Blair, and Carol Stocking, "Modest Expectations: The Effects of Interviewers' Prior Expectations on Response," *Sociological Methods and Research*, Vol. 6, 1977, pp. 171-182.

Tepping, Benjamin J., and Barbara A. Bailar, "Enumerator Variance in the 1970 Census," *Proceedings of the American Statistical Association, Social Statistics Section*, 1973.