Economic and Social Review Vol. 9 No. 3

Some Evidence on the Validity of Survey Results

B. J. WHELAN
The Economic and Social Research Institute, Dublin

C. A. Ó MUIRCHEARTAIGH
London School of Economics

I INTRODUCTION

Data from sample surveys are subject to many sources of error. The most commonly discussed of these are *sampling* errors i.e., errors in the estimates which arise as a result of not including the whole population in the enumeration process. However, even when a complete enumeration of the population is carried out, the data obtained (and the conclusions reached) may be subject to serious error due to faults in the method of measurement or observation. These response errors are a component of the total error in the survey estimates and may arise from faults in the questionnaire, from badly conducted fieldwork, or from inaccuracy in the respondent's answers.

The error in an individual response may be defined as the difference between the response recorded and the individual true value (itv). This itv must be independent of the conditions under which the survey is carried out. Hansen, Hurwitz and Madow (1953) defined the itv as follows:

(a) it must be uniquely defined;
(b) it must be defined in such a manner that the purposes of the survey are met;
(c) where it is possible to do so consistently with the first two criteria, it should be defined in terms of operations that can actually be carried out (even though it might be difficult or expensive to do so). In certain cases, even the definition of the itv can cause problems. With factual

variables such as age or sex, the problems of definition are not severe, but with attitudinal variables these problems can be considerable.

The individual response error may be divided into two components: the *individual response bias* (i.e., the constant or systematic deviation from the true value) and the remainder which is called the *response deviation.* The *response bias* is the average of the individual response biases over the whole population.

Ideally, the researcher would like to be able to estimate the total error from all sources in his data. Though eminently desirable, this is rarely possible. One can usually estimate variable sampling errors (sampling variance) since there exists a well developed theory to deal with them. (See, for instance, Kish (1965) or Cochran (1977).) However, relatively little is known about the magnitude, determinants and effects of non-sampling errors. Some investigations have been carried out into variable non-sampling errors such as interviewer variance (e.g., Kish (1962); Fellegi (1964, 1974); Hansen, Hurwitz and Bershad (1961); Ó Muircheartaigh (1976).) The difficulty with such studies is that modifications must be built into the survey design to accommodate them. Even less is known about constant non-sampling errors. Apart from some early studies (e.g. Parry and Crossley (1950)), the bulk of the most interesting findings have been obtained by Cannell and his colleagues in a series of experimental studies (Cannell *et al.* (1970), Cannell (1977)).

The explanation of the paucity of results in this area is two-fold. First, some non-sampling errors tend to be specific to the survey in question and it is very difficult to derive general principles about their operation. Furthermore, it is usually impossible, or prohibitively expensive, to obtain data about the "individual true values" (itv's) Indeed, as was pointed out above, even the definition of what constitutes the true value can be controversial. Close approximation to the itv is necessary if one is to calculate the gross error for a certain variable. It is this necessity to have a reliable and valid measure of the individual value which makes the estimation of the constant non-sampling errors so difficult outside experimental studies.

The sample examined in the present study provided a unique opportunity to estimate the overall error in measuring certain key variables. The subject of inquiry was redundancy – who are the redundant, how do they cope after redundancy, how many obtain new jobs, etc? The substantive results are written up in Whelan and Walsh (1977). Interviewing was carried out in late 1974 about 2½ years after the respondents were made redundant. The sampling frame consisted of the records of the Irish Department of Labour and the Department of Social Welfare. These contained detailed information on each redundant worker (age, sex, pay before redundancy, size of lump sum, length of service, etc.). A stratified sample of 602 workers was selected

from these files and interviews were attempted with each worker. In all, 503 interviews were obtained. The interviewers involved were professionals from the field force of The Economic and Social Research Institute.

This study tries to assess the magnitude and determinants of response errors (i.e. deviation of the answer recorded for the respondent from its "true" value) for four important variables. These deviations are the net error from all sources for each respondent. We are, therefore, concerned only with that portion of the total error which is due to deviations from the itv and not with the portion attributable to sampling variance or sampling bias. Two of the variables in question – age and income (in our case, income before redundancy) –are utilised in almost every social survey and the accuracy with which they can be measured is of crucial importance. The other two were important variables in the study of redundancy: the redundancy payment received and the proportion of the two years following redundancy which the respondent spent in employment. It is hoped that our findings will provide some evidence on the accuracy of these types of survey data, and suggest how improvements in accuracy might be attained.

II SOURCES OF VALIDATION DATA

Before going on to analyse the response errors, a few words are necessary concerning the sources of our data on the "true" values. Age, pay before redundancy, and lump sum received were obtained from the Department of Labour records which we used as a sampling frame. We have a high degree of confidence in the quality of these data since each of these items was of importance in calculating the worker's entitlement to a redundancy lump sum and redundancy pay. The other "true" value (time employed after redundancy) was computed from social insurance records and is probably less valid than the others due to problems of definition and measurement. However, we are, on the whole, satisfied that the four "true" values are more valid measures than are the corresponding survey items. Indeed, in many studies of this kind, the validation data (or estimated itv's) are based on subjective expert assessment and not on objective criteria. For example, in Parry and Crossley (1950), much of the validation data was obtained from municipal records; Kish and Lansing (1954) used estate agents' (realtors') estimates of the values of houses.

III RESULTS

We first of all present the estimated response biases for the four variables. These are shown in Table 1. It is encouraging to note that the biases seem

Table 1: *Survey estimates and true values for the four variables*

Variable	*Survey estimate*	*True value*	*Percentage error*
Pre-redundancy pay (£ per week)	22.3	21.9	1.75
Redundancy payment (£)	251.0	201.2	24.75
Percentage of time since redundancy spent in employment (%)	62.2	60.3	3.15
Age (years)	42.6	42.4	0.40

very low for Items 1 (Pay), 3 (Employment), and 4 (Age), all three being less than 4 per cent of the true value. The bias in the reported State redundancy payment is much larger being about 24 per cent of the true value. Several factors are, undoubtedly, responsible for this large positive bias: (i) the payment was made over two years before the date of interview and memory bias may have crept in; (ii) many respondents received income from several sources at redundancy (golden handshake, back pay, etc.) and may, therefore, never have known the exact amount of the State redundancy payment.[1]

If we are to go beyond the response biases and investigate the sources of variation in the individual response errors, it is necessary to analyse the response deviations and seek to identify their causes, or correlates. One possible hypothesis is that, in addition to the overall bias, differences in interviewer quality contribute to the total error (e.g. more able interviewers get more accurate data). Another is that respondent characteristics have some effect (e.g. older respondents are more prone to lapses of memory). Tables 2 and 3 present some tests for the presence of such effects.

Table 2 reports on an Analysis of Variance of the response deviations by interviewer. Only the State redundancy payment showed a significant interviewer effect. It should be noted that the absence of randomisation of respondents to interviewers may have affected the results of these tests. The use of stratified rather than simple random sampling may also have affected the efficiency of the tests, to the extent that the resulting error distributions were non-normal.

Table 3 presents the results of four regressions of the response deviations on respondent characteristics and interviewer characteristics. The first point to note about this table is the small number of significant coefficients and the low values of R^2. This means that the variables listed explain only a very small proportion of the variance in the response deviations. The regression

[1] It might be noted, however, that the questionnaire used indicated clearly that it was the "State Redundancy Payment" which was required.

Table 2: *Analysis of variance by interviewer.*

	Dependent variables: Response deviations in			
	Pre-redundancy pay (£)	*State redundancy payment (£)*	*Time employed (% of 2 years)*	*Age (years)*
F =	1.20	1.56	1.38	0.41
d.f. =	(29,370)	(30,318)	(30,265)	(30,463)

of the deviation in the State redundancy payment contains only one significant coefficient, that of the "true value". This probably reflects a tendency for those who received higher payments to round upwards. The response deviation in the employment variable has a significant relationship with only two variables: the true value and respondent's sex. The difference between

Table 3: *Regression results*

	Dependent variables: Response deviations in			
	Pre-redundancy pay (£)	*State redundancy payment (£)*	*Time employed (% of 2 years)*	*Age (years)*
Independent Variables	*Coefficients*			
True pre-redundancy pay	−0.27**	0.74	0.18	−0.03
True redundancy payment	−0.00	0.18*	−.001	0.00
True time employed	0.99	−12.15	17.55**	−1.09
True age	−0.22	−1.57	0.02	−0.07**
Sex (Male = 1)	2.88**	12.50	7.69*	–
Years education	0.22	−4.91	0.56	−0.02
Rural residence	−0.64	−27.81	1.17	0.49
Interviewer's rating	0.24	−9.94	0.41	0.04
Interviewer's experience	−0.23	0.68	−0.31	–
Interviewer's non-response rate	0.03	1.42	0.06	0.01
Constant	1.80	153.33	3.30	3.62
R^2	0.16	0.06	0.10	0.04

* Significant at 5% level.

** Significant at 1% level.

the sexes may be explained by the fact that the social insurance records define "unemployment" as being in receipt of benefits, while the survey defined it in terms of willingness to take a job. On the whole, therefore, our survey tended to turn up more unemployed women that appeared from the social welfare records. Only true age is significant in the regression of the deviation in age. This result is easy to rationalise on the assumption that older people tend to understate their age.

In view of the importance of income data for many surveys, and also because of the relatively high R^2, the equation of the deviation in pre-redundancy pay is very interesting. It implies that women and those on higher incomes are likely to understate their income. From the respondent's point of view, such understatement is quite rational if he believes that the results of the inquiry may lead to his being deprived of social welfare benefits or taxed more heavily. Beliefs of this sort are quite often expressed by respondents in interviews. Our findings on this point underline the importance of the interviewer's initial approach to the respondent and skill in convincing him of the complete confidentiality of his answers.

A major feature of all four equations is that the interviewer characteristics are never significant. This corroborates the comparative lack of interviewer effects noted in the ANOVA results in Table 2.

IV CONCLUSIONS

On the whole, this study gives grounds for confidence in the validity of the results of the survey. Of the four variables examined, only the State redundancy payment showed any appreciable bias. This was also the only variable which had a significant interviewer effect. Regression analysis of the response deviations did not reveal any strong pattern except in the case of pre-redundancy pay. The results for this variable suggest that male respondents tend to over-rate their pay, while those on high pay tend to understate it. Interviewer characteristics had no discernible influence on any of the response deviations studied.

While the results of this analysis demonstrate that surveys can provide valid and accurate results, there are several reasons why some caution is to be recommended before extrapolating this finding to the results of all surveys. Firstly, the content of the questionnaire was almost entirely *factual* and previous research has shown that opinion questions are more sensitive to interviewer effects than are factual questions. Secondly, the above analysis has concentrated on overall averages. The effects of response deviations may be more marked in certain sub-groups in the population. Finally, it

should be borne in mind that the interviewing was carried out by a team of carefully briefed, experienced interviewers and that the questionnaire had been fully pre-tested. Less valid results would undoubtedly be obtained from surveys where these conditions did not prevail.

REFERENCES

CANNELL, C. F., K. H. MARQUIS, and A. LAURENT 1970. *A Summary of Research Studies in Interviewing Methodology*. Ann Arbour: Survey Research Centre, University of Michigan.

CANNELL, C. F. 1977. *Interviewing Methodology*, Washington, D. C.: National Centre for Health Statistics. US Government Printing Office.

COCHRAN, W. G. 1977. *Sampling Techniques*, New York: John Wiley & Sons Inc.

FELLEGI, I. P. 1964. "Response variance and its estimation" *Journal of the American Statistical Association*, Vol. 59, pp. 1016–1041.

FELLEGI, I. P. 1974. "An improved method of estimating the correlated response variance" *Journal of the American Statistical Association*, Vol. 69, pp. 496–501.

HANSEN, M. H., W. N. HURWITZ, and M. A. BERSHAD, 1961. "Measurement of errors in censuses and surveys" *Bulletin of the International Statistical Institute*, Vol. 38, pp. 359–374.

HANSEN, M. H., W. N. HURWITZ, and W. G. MADOW, 1953. *Sample Survey Methods and Theory*, New York: John Wiley & Sons Inc.

KISH, L. 1962. "Studies of interviewer variance for attitudinal variables" *Journal of the American Statistical Association*, Vol. 57, pp. 92–115.

KISH, L. 1965. *Survey Sampling*, New York: John Wiley & Sons Inc.

KISH, L. and J. B. LANSING, 1954. "Response errors in estimating the value of homes" *Journal of the American Statistical Association*, Vol. 49, pp. 520–538.

Ó MUIRCHEARTAIGH, C. A. 1976. "Response Errors in an attitudinal sample survey" *Quality & Quantity* Vol. 10. No. 1. pp. 97–115.

PARRY, H. J. and H. M. CROSSLEY, 1950. "Validity of responses to survey questions" *Public Opinion Quarterly*, Vol. 14, pp. 61–80.

WHELAN, B. J. and B. M. WALSH, 1977 *Redundancy and Re-employment in Ireland*, Dublin: The Economic and Social Research Institute, Paper No. 89.