RESPONDENT-GENERATED INTERVAL ESTIMATION IN SAMPLE SURVEYS MAY REDUCE ITEM-NONRESPONSE[1]

S. James Press, University of California at Riverside; Judith M. Tanur, State University of New York at Stony Brook

Judith M. Tanur, Department of Sociology, State University of New York at Stony Brook, Stony Brook, NY 11794-4356

Key Words: Recall, Nonresponse, Intervals

Abstract

Press (1999) proposed a method of questioning in surveys expected to increase the accuracy of estimators of population means obtained by recall from respondents. The questioning procedure is called Respondent-Generated Intervals (RGI). The method involves requesting that a respondent supply not only a point assessment of a quantity, but also an assessment of the smallest and largest values the quantity could possibly be. Point and interval estimates of the population mean are then generated based on these assessments. Press and Tanur (1999) show that these estimates are quite accurate. Depending on the item, between 1% and 14% of respondents supplied an interval in which they believed the quantity was almost certain to fall, but seemed not to be sure enough of the value of the quantity to answer the direct question. Perhaps these respondents would have simply skipped the question if they had not been permitted to offer an interval. If that is the case, we find savings in item non-response achieved by the offering of such intervals. Perhaps asking for an interval *instead* of asking a direct question (rather than asking for both) will reduce item nonresponse. For some of the cases in which only intervals were offered we have verification data, and so are also able to examine the accuracy achieved by using the midpoint of the respondent-generated interval as an estimate of the quantity of interest. We also find that the intervals formed by the averages of the endpoints of the respondent-generated intervals cover the true values of the quantities being estimated for all cases we can examine.

Survey researchers do not need to be reminded of the importance of the problem of nonresponse. Methods for avoiding it, minimizing it, measuring it, and compensating for it when it have generated a large literature (e.g. Panel on Incomplete Data, 1983; Groves,1989). This paper presents a simple method which seems to have some promise for reducing item nonresponse, a method that presented itself to us as a by-product of research we have been conducting in an attempt to improve recall accuracy. That method asks respondents to give an interval that they believe to bracket the quantitative item they are being asked to recall.

Of course, intervals have long been used in survey questionnaires, notably in questions about income, where it is believed that respondents are more willing to report on the sensitive issue of income within a range rather than with an exact number. But the intervals used in the method presented here are generated by the respondent rather than by the survey researcher, which is why we call them Respondent-Generated Intervals, RGI. In an experiment we ran on two college campuses to study the characteristics of the RGI protocol, we noticed that a sizable number of respondents supplied an interval in which they believed the quantity was almost certain to fall, but seemed not to be sure enough of the value of the quantity to answer the direct question. Perhaps these respondents would have simply skipped the question if they had not been permitted to offer an interval. On the assumption that such respondents would have indeed been item nonrespondents, in this paper we explore the savings in item non-response that could be achieved by the offering of such intervals.

[1]We are very grateful to Dr. Robert Creecy for his suggestions and computer programming help; to Mr. Jacob Pozharny, Ms Lucinda Manolakes, Mr. Andrew Catalano, Mr. Nephtali O'Connor, Mr. Alex Trillo, Ms. Warangkana Sangchan, and Mr. Jeffery Dennis for their help in collecting, cleaning and analyzing the data; to Dr. Kent Marquis for helpful discussions and suggestions; to the U.S. Census Bureau for its financial support; and to the campus administrators who provided verification.

Data and Methods

At our respective universities we administered surveys designed to test a method, proposed by Press (see Press, 1999), expected to increase the accuracy of quantitative data recalled by respondents. The method involves requesting that a respondent supply not only a point assessment of the quantity s/he is asked to recall (what we have called the "usage quantity"), but also an interval assessment of the smallest and largest values s/he believes the quantity could possibly be (RGI). Point and interval estimates of the population mean were then generated based on these assessments. Results of these analyses appear in Press and Tanur (1999), and seem very promising.

These data were generated in the spring of 1997 when approximately 1000 students at the University of California at Riverside (UCR) and about 750 at the State University of New York at Stony Brook (SUSB) responded to a paper-and-pencil, self-administered questionnaire asking them to recall factual information about matters relating to their life on campus. At both campuses students were asked for the number of credits they had earned (CRED), the number of grades of C or less they had received (C's), their grade point average (GPA), their Scholastic Aptitude Scores on the math (SATM) and verbal (SATV) tests[2], and the number of traffic tickets they had received on campus during that academic year (TICK). At UCR students were also asked to recall the amount of the registration fee (REG) and the recreation center fee (REC) they paid at the beginning of the quarter. The corresponding questions at Stony Brook asked for the amounts of the health fee (HFEE) and the student activities fee (SAFEE). Stony Brook students were also asked to recall the amount spent on the food plan (FOOD) and the number of library fines (FINES) they had been assessed. As stated above, we call the answers to these recall questions "usage quantities."

For each usage quantity, respondents were also asked to give an interval in which they were almost certain that the quantity would lie. (In fact, because one of our aims was to test two possible ways of asking this "uncertainty" question, we used a split ballot design, so only half of each sample was offered the interval question as here described for most recall items. For the fee data, however, all respondents were presented with the interval question.) For those respondents who consented to have their data verified and who gave their student identification numbers for that purpose, we were able to get "true" values for the usage quantities the students were recalling from appropriate campus offices. (No permission for verification or identification numbers were needed to ascertain the true values for the fee data, as the fees are standard for all full-time students. We limited our analysis to full-time students.) While we are cognizant of the possible errors in the administrative data we are using for verification, we shall use those data as a "gold standard" in what follows.

[2] The SAT verbal and math scores are not used in this analysis because we are unable to separate data missing because the respondent did not take the SAT tests from data missing for reasons that the interval procedure can hope to remedy.

Findings

Table 1 shows the number of responses for each question classified as full responses (both a usage quantity and an interval were given, column 3), usage but no interval (column 1), interval but no usage (column 2), and neither usage nor interval (column 4). Column 5 shows the total number of potential respondents, column 6 presents the interval-only responses as a percentage of the total potential respondents to the question (i.e., column 2/column 5), and column 7 presents the interval-only respondents as a percentage of nonresponders to the usage quantity (i.e., column 2/[column 2 + column 4]). We see from column 6 that those who chose to give an interval-only response constitute only a few percent of the total potential responders for all questions except for the fee questions at SUSB where those percents are 9.7% and 13.4%. But we see from column 7 that the interval-only responders constitute an appreciable percentage of potential nonresponders to each item. Indeed, those percents are never less than 4% and twice are over 40%. We can interpret these results as estimated conditional probabilities of giving an interval among those who did not give a usage quantity.

Of course, having ascertained that respectable proportions of respondents who do not answer the direct question do provide intervals, we must inquire about the accuracy of such intervals if they should be used to produce estimators of the usage quantities. Press (1999) proposed (among other estimators) that the average midpoint of the interval be used as a point estimate of the usage quantity and that the interval formed by the average of the lower bounds supplied by the respondents and the average of their upper bounds (termed the Average Respondent-Generated Interval, ARGI) be used as an interval estimator. We shall evaluate the accuracy of these estimators for those respondents who offered interval but no usage quantity responses.

Unfortunately, at UCR, no verification data were collected for respondents who omitted a usage quantity response, and at SUSB the number of respondents for which such validation data are available is tiny (n ranges from 0 to 5). Fortunately, however, verification data are not needed for the fee questions, for which the values are constant for all full-time students, so we can use the fee data to evaluate the accuracy of the estimators. From Table 2 we can see that the average midpoints of the RGIs overestimate the true usage quantities in three out of four cases. The ARGIs, however, cover that true value in all four cases.

The results so far, however, lump together two kinds of respondents. Some students participate in funding their college education, and hence the task of reporting on the amount paid for fees is almost purely one of recall from autobiographical memory. Other students, however, do not so participate; for them, the reporting task may well include elements of proxy reporting for amounts mentioned to them by their parents or noticed on financial aid forms. It seemed sensible to look at the results shown in Table 2 separately for those students who participated in funding their education and those who did not. These results are shown in Table 3.

Table 3 shows that the midpoint of the RGI is always closer to truth for the self-payers than it is for the non-self payers. The ARGI, however, covers truth only for two of the four items for the self-payers and three of the four for the non-self-payers. The final two columns of Table 3 present another interval estimator proposed by Press (1999), the ARGI (1σ), which adds one estimated standard error of the upper bounds to the average upper bound and subtracts one estimated standard error of the lower bounds from the average lower bound. Examining the ARGI (1σ), we find that the extended interval covers truth for all items except the UCR registration fee for both self-payers and non-self-payers.

We predicate our thinking about respondents who gave an interval but no usage quantity on the premise that they are trying their utmost to be cooperative. Indeed, we visualize individuals sufficiently unsure of the correct answer that they do not wish to risk giving a point estimate, but taking the survey task sufficiently seriously that they are willing to codify their partial knowledge into an interval. The fee data, for which there are an appreciable number of RGI-only respondents, afford us the opportunity for a partial test of this premise. In addition, the fee data recommend themselves for this analysis because these questions seem to embody few elements of threat or of social desirability. Hence they should not pose special problems to challenge or to abet the cooperation of respondents.

To test the premise that "interval only" respondents are cooperators we ran logistic regressions to predict the probability of giving an interval when a usage quantity was not given. We asked whether those who we can conceptualize as more prone to cooperation are more likely to offer an interval than those who we cannot conceptualize as prone to cooperation. We operationalize "proneness to cooperation" with three independent variables: whether a respondent gave a social security number, thus granting permission for verification of his/her data (SSNIND), year in school (as a rough measure of maturity -- YEAR), and GPA (as a rough indicator of "playing by the rules" at the university). In the logistic regression we used whether the respondent supplied an interval as the dependent variable; we also introduce an indicator of whether the respondent participated in funding his/her education (SELFPAY), reasoning as above that those who do so participate will be more cognizant of the amount of fees paid.

The results of the logistic regressions are shown in Table 4. We see that the analyses for both fees at UCR show very poor model fit. We speculate that this is a result of the relatively small sample sizes for the UCR fee data (relative to the SUSB fee data) and the fact that 20% or fewer of the UCR no-usage-quantity respondents offered an interval. The smaller sample sizes may have resulted in insufficient power of the Wald test to detect significance of the relevant coefficients. At SUSB, however, in both cases over 40% of the no-usage-quantity respondents supplied an interval, and the results are more revealing. Both logistic regression models at SUSB exhibit statistically significant fits, and the coefficients of all the independent variables are positive, indicating that those respondents whom we would expect to be more cooperative (or more knowledgeable) are indeed more likely to supply an interval. The social security number indicator (indicating the respondent's willingness to have his/her data verified from academic records) attains statistical significance at the .044 level for the student activities fee, and year in school (YEAR) and SELFPAY are statistically significant for the health fee at $p=.076$ and $p=.066$ respectively. Thus those respondents we would expect to be more prone to cooperate do seem to cooperate more often.

Conclusions

We have shown that, in a substantial proportion of cases, respondents who do not supply an estimate of usage quantities do indeed supply intervals, thus reducing the amount of item nonresponse

appreciably. Further, in all cases we could examine (the four fee items), the ARGI covered the true value of the usage quantity, thus supplying useful data for respondents who would otherwise have been coded as missing. Those who participated in funding their own education gave slightly more accurate intervals when they did not give a usage quantity than those who did not so participate, and we have some evidence that those respondents we would expect to cooperate are more likely to give an interval when they do not give a usage quantity than are those who we would not expect to cooperate. Of course, the reader should note that these respondents are already in some sense uncooperative, having refused to follow instructions to give both a usage quantity and an interval. Thus, although whether asking a respondent to supply only an interval will be effective as a device to reduce item nonresponse cannot be tested with these data, the results are suggestive and merit further research

References

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley

Panel on Incomplete Data 1983 *Incomplete Data in Sample Surveys.* New York: Academic Press.

Press, S. James and Tanur, Judith (1999) "Experimenting with Respondent-Generated Intervals", manuscript submitted for publication, and available for viewing and downloading at the website http://cnas.ucr.edu/~stat/press.htm

Press, S. James (1999). "Respondent-Generated Intervals For Recall in Sample Surveys", manuscript submitted for publication, and available for viewing and downloading at the website http://cnas.ucr.edu/~stat/press.htm

Table 1: Responses Classified by Completeness/incompleteness

Item	1 Usage, no Interval	2 Interval, no Usage	3 Interval and Usage	4 Neither Interval nor Usage	5 Total Respondents	6 Percent of Total giving Interval but Not Usage	7 Percent of "Nonresponders" giving Interval but Not Usage
SUSB							
Credits	39	4	328	11	382	1.0%	26.7%
C's	41	11	309	16	377	2.9%	40.7%
Fines	59	2	282	40	383	0.5%	4.8%
Food	7	7	158	35	207	3.4%	16.7%
GPA	26	2	340	15	383	0.5%	11.8%
Health Fee	27	74	509	150	760	9.7%	33.0%
SAFee	29	102	483	146	760	13.4%	41.1%
Tickets	50	6	294	27	377	1.6%	18.2%
UCR							
Credits	38	8	437	25	508	1.6%	24.2%
C's	41	6	447	24	518	1.2%	20.0%
GPA	35	4	444	25	508	0.8%	13.8%
RecFee	52	26	827	121	1026	2.5%	17.7%
RegFee	49	24	779	174	1026	2.3%	12.1%
Tickets	31	2	462	23	518	0.4%	8.0%

Table 2: Accuracy of Respondent-Generated Intervals for RGI-only Respondents: Fee Data

Item	*N*	Truth	Average Midpoint	Av. Lower Bound	Av. Upper Bound
SUSB Health Fee	74	$70.00	$123.98	$41.92	$206.04
SUSB Activity Fee	102	$78.75	$87.14	$33.06	$141.23
UCR Rec Fee	26	$59.00	$121.71	$47.73	$195.69
UCR Reg Fee	24	$1,371.25	$1,343.65	$922.33	$1,764.96

Table 3: Accuracy of Respondent-Generated Intervals for RGI-only Respondents by Payment Status: Fee Data

			AVERAGE RGI			AVERAGE RGI (1s)	
			AVERAGE	*AVERAGE LOWER*	*AVERAGE UPPER*	*AVERAGE LOWER*	*AVERAGE UPPER*
	TRUTH	*N*	*MIDPOINT*	*BOUND*	*BOUND*	*BOUND - 1s*	*BOUND + 1s*
ITEM					**SELF-PAYERS**		
SUSB Health Fee	$70.00	39	$120.14	$47.31	$192.97	$39.26	$224.77
SUSB Activity Fee	$78.75	48	$98.90	$36.65	$161.15	$30.69	$188.37
UCR Rec Fee	$59.00	6	$33.08	$20.17	$46.00	$12.06	$61.32
UCR Reg Fee	$1,371.25	8	$895.35	$857.38	$933.33	$729.01	$1,123.15
					NON-SELF-PAYERS		
SUSB Health Fee	$70.00	35	$128.26	$35.91	$220.60	$29.34	$265.73
SUSB Activity Fee	$78.75	49	$117.38	$48.44	$186.33	$28.11	$248.52
UCR Rec Fee	$59.00	18	$155.53	$54.72	$256.33	$38.36	$316.84
UCR Reg Fee	$1,371.25	18	$734.36	$613.70	$855.56	$465.30	$1,098.18

Table 4: Coefficients of logistic regressions--Fee data

(p values given for results statistically significant at .10 or below)

	UCR REG	UCR REC	SUSB ACTIV.	SUSB HEALTH
SSNIND	7.9195	--	0.651 *(p=.044)*	0.5472
YEAR	-0.1165	0.2625	0.1292	0.2493 *(p=.076)*
GPA	-0.7029 *(p=.097)*	-0.4326	0.4077	0.2579
SELFPAY	0.4412	0.037	0.4195	0.5994 *(p=.066)*
CONSTANT	0.7624	-0.6907	-2.1302 *(p=.012)*	-2.2191 *(p=.015)*
Model Chi Square (df)	6.502 (4)	2.837 (3)	9.925 (4) *(p=.042)*	11.110 (4) *(p=.042)*
Percent giving RGI	18.31%	20.19%	45.90%	41.80%
N	142	104	183	172