

NONRESPONSE FOLLOW-UP IN ESTABLISHMENT SURVEYS: A SPLIT-HALF EXPERIMENT

Young I. Chun, Bureau of Labor Statistics
2 Massachusetts Avenue, N.E. Washington, DC 20212

Key words: Nonresponse, Establishment survey, and Experiment

1. INTRODUCTION

Follow-ups in establishment surveys are persuasive efforts to reduce the nonresponse rate by converting nonresponse. Efforts to convert nonrespondents to an establishment survey include a reminder/thank-you letter, telephone or FAX prompting, additional mailings of replacement questionnaires by first class or certified mail, and face-to-face interview or telephone interview (Dillman, 1978; Groves, 1989; House et al., 1976). The success of follow-ups are important as higher nonresponse rates have remained a major drawback of establishment surveys. Nonresponse has two main effects on measurement errors (Lyberg and Dean, 1992). Nonresponse increases variance since the collected sample size is smaller than the intended sample size, and this difference may vary over conceptual replications of a survey. Nonresponse also may increase bias that is a function of the nonresponse rate and the differences in the sample statistics between respondents and nonrespondents.

The establishment survey discussed in this paper focuses on the establishment as the unit of analysis. The informant's reporting task at an establishment primarily depends on the retrieval of records from an organization's information system. The purpose of this paper is to evaluate in an experimental study a nonresponse telephone prompting follow-up approach against a technique currently employed in the Hours at Work Survey (HWS), which is conducted by the Bureau of Labor Statistics (BLS).

Nonrespondents in the experiment received the two follow-up mailings before being contacted by telephone. The current nonresponse follow-up approach uses a telephone interview which asks the potential respondent to provide an *estimate* of hours paid and hours at work. The new telephone-prompt approach requires interviewers to identify and trace an appropriate informant, and asks the informant to use the establishment records based *hard data*. The informant is asked to return the completed form by fax. The interviewer is instructed to take the industry type and size of the nonresponding establishment into consideration when establishing a rapport with the informant. The experiment was designed to address

three key issues in follow-ups: data quality, cost, and response rate.

2. LITERATURE REVIEW AND HYPOTHESES

Studies on the effect of follow-ups in establishment surveys are very limited and their focus has been on recording response improvement only. Over the last few decades, survey researchers have utilized the Total Design Method (TDM) and have used certified-mail as a last attempt to convert strong "hold-outs." Dillman (1978) reports that the certified mail follow-up raises response rates 13 percentage points on the average and produces a greater relative return than any follow-up mailing that preceded it. These findings are based on household surveys. In implementing the TDM in the early stage of establishment survey data collection, Chun and Robertson (1996) found in their experimental study that the test group receiving advance and reminder/thank you letters increased the response rate by 7.5 percentage points, when compared to the control group receiving neither letters.

In the Current Employment Statistics Survey of over 380,000 business establishments monthly, Rosen et al. (1992) studied the use of FAX communication as a substitute for the advance notice postcard and the nonresponse telephone prompt call. They concluded that the wide availability of FAX machines makes the use of FAX a viable, convenient, and cost-effective option for nonrespondent conversion in an establishment survey. With the exception of those establishments with less than 5 employees, more than 75% of establishments possess a FAX machine. They found the combined effect of FAX advance notice and FAX nonresponse prompting yields essentially the same response rate as the current postcard notice and telephone prompting to nonrespondents. Furthermore, they found FAX offers a potential cost advantage over postcard prenotice and telephone prompting. Data quality was not examined in this study.

Research also shows that the effect of follow-ups/repeated contacts on response rates is moderated by the type of sample (i.e., consumer vs. institutional groups). Yammarino and his colleagues (1991) suggest in their meta-analysis of mail surveys that, among institutional groups, the correlation of follow-up/repeated contacts with response rate was four times higher than among consumer groups. The variance explained by sampling error in the institutional groups

was .54, as compared to .14 in the consumer groups. Overall, follow-ups for institutional groups increase response rate by 30.6 percent. The type of sample appeared to moderate the effect of the follow-ups on response rate.

Despite the very significant effects of general follow-ups, research fails to document the effects of follow-ups on the difficult cases in late stage of establishment survey data collection. The experimental study discussed in this paper was conducted where half of the nonrespondents remaining after a second follow-up mailing were randomly selected to be followed by asking them to provide an estimate of hours paid and hours at work, and the remaining half were prompted by telephone to fill out the questionnaire with the hard data developed from the organization's information system. This approach is an answer to Lyberg and Dean (1992) who called for sequential designs for reducing 'hard-core holdouts.' They questioned the validity of the current studies for nonresponse because of a tendency to follow-up on the easiest cases.

The current follow-up approach used after the second mail follow-up is designed to ask for estimates of hours paid and hours at work. Respondents were encouraged to make educated estimates of the hours information based on the number of employees and the average leave hours. In contrast, the alternative telephone prompting is developed to identify a most appropriate respondent, confirm the respondent is knowledgeable about business records to be used, and ask the respondent to fill out the original mail survey form based on company records. Interviewers using nonresponse prompting were intensively trained to learn basic call back rules, approaches to establish a good rapport, and persuasive arguments to reply to various reasons for refusals.

Three hypotheses are considered regarding the effects of the above follow-ups on data quality, cost-effectiveness, and nonresponse reduction. The hypothesis regarding the issue of data quality is not evaluated here as it is fairly obvious and well documented previously. A response analysis survey found that data from the original mail survey instrument is more complete and accurate than the estimates that respondents provide during a brief telephone call. Therefore, data collected with telephone prompting is of higher quality than data collected by the current telephone follow-up.

Regarding cost-effectiveness, it is hypothesized that hours spent on the telephone prompting would be close to those on the telephone estimate. In the telephone prompting, interviewers spend much of their time to

locate and identify a correct contact. In the telephone estimate, interviewers spend most of their time, once they speak with a reasonable contact, to get estimates on various questions. We expect the time would be about the same across two follow-up methods.

In terms of response rate, it is expected that the telephone prompting would not reduce nonresponse rate as much as would the telephone estimate. Respondents to the telephone prompting tend to rely on their record-keeping system to extract the data required, and discuss with their associates to provide reliable data. As a result, they provide quality data at the expense of low compliance. In contrast, the nonresponse reduction by the telephone estimate is likely to be relatively higher than the telephone prompting. Estimates are easy to make once compliance is earned. The overall nonresponse reduction is thus likely to be high with the telephone estimate.

3. RESEARCH DESIGN

The experiment ($n = 1,200$) was embedded in the HWS which collected data for the 1995 reference period. A probability sample of nonrespondents to the initial mail survey was selected by stratifying by two major industrial divisions and by four employment size classes. Half of the sample was randomly assigned to the telephone estimate group; the remaining half, to the telephone prompting group. In the telephone estimate group, "hard-core" holdouts were encouraged to provide their educated estimates based on rough figures of employment and average leave hours. In the telephone prompting group, nonrespondents were contacted by interviewers who were intensively trained regarding methods to build a rapport and persuasive arguments to answer all different reasons for refusals. We prepared a training agenda, drawing on insights from nonresponse conversion efforts in telephone follow-up surveys. The telephone prompter training included practice of scripted telephone prompting procedures including appropriate reactions to various reasons for refusal, discussion of persuasive techniques, and use of call record sheets. Approaches to locate the sample subject and find the best time to call back were also included in the training. Interviewers conducted mock-up surveys, and were monitored regarding their interactions with respondents.

The HWS is a national annual survey of 6,000 establishments conducted by the BLS since 1981. The HWS primarily collects data on both the total number of hours which nonsupervisory or production employees were on the job (i.e., hours paid minus paid leave) and the total number of hours for which these employees were paid. The information obtained is used

as one factor in the estimation of national productivity by industry.

4. FINDINGS AND DISCUSSION

It is important to note that findings reported here should be interpreted with due caution as there are several confounding factors which would make a complete analysis of this experiment difficult. The two groups were collected over different length time frames. We learned from previous tests that a longer data collection period lowers response per unit of time at some point. Second, there were different levels of incompleteness in this experiment due to the lack of resources being applied to either group. Finally, there is an unequal application of human resources across the test and control groups.

Cost-effectiveness: Results indicate that the prompting-script based method converts more nonresponding units than the current method *per hour*. Table 1 indicates that the number of usable units collected per 100 hours of interviewer time is greater for the telephone prompting approach (31.6 usable units) than for the current telephone estimate (24.9 usable units). Although this is not a significant increase in response as indicated by the t-statistics ($\alpha = 0.05$), it is very promising as we were expecting a decrease in response. When we look at the average number of hours needed to obtain a usable response, it is obvious that the telephone prompting required less interviewer time on average to obtain a response than the telephone estimate did (3.17 vs. 4.01 hours per usable response). It is noteworthy that the telephone prompting group was collected across a longer time-frame than the estimate group. It is suspected from previous findings that more resources applied quicker provides a better overall response. Due to resource limitations, the test was not completed as planned. Therefore, while the results are encouraging, they cannot be accepted as defining what a complete test would yield with enough resources.

Response Rates: Our expectations for response results were modest. The response burden for the telephone prompting is significantly higher than the burden for the telephone estimates. Once a respondent is reached, the telephone estimate takes only a few minute to elicit a final response about hours paid and hours at work. In contrast, the telephone prompting asks them to look after their records, fill the original mail questionnaire out, and fax it to us. We learned from a small pilot test that the majority of potential respondents in telephone follow-ups had never actually seen the survey questionnaire yet. Therefore, we did not want to assume that the contact would be unwilling to devote the time to providing hard data developed from the information system. We also learned from previous

small tests that the time required to prompt a respondent to provide either a faxed report or an estimate is small compared to the time devoted to actually identifying who that respondent is at this point in the months-long survey.

Table 2 shows that the telephone prompting group had a final response rate 18.7 percentage points lower than the telephone estimate group, and this difference is significant at $\alpha = 0.05$ level. An initial reaction might be that this is an unfavorable finding, if we do not take into consideration the time spent on data collection for these two groups. As shown in the Table 2, the total time spent on the telephone prompting was 380 hours, while the time spent on the telephone estimate was over two times as much. Had equal resources been applied to the two groups, we would expect the response rates to be considerably closer together.

Table 3 lays out the distribution of units into various response dispositions for both groups. As expected, there are more refusals in the telephone prompting group than in the telephone estimate group. The telephone prompting approach encourages the respondent to provide with the record-based hard data, and there is a high reluctance to follow on the respondent side. With the telephone prompting approach, interviewers confirmed the status of all units except only 2 %, and found relatively greater number of units not locatable. The 17 units that were tentatively defined with unusable data may have been made usable if there had been sufficient resources to follow up on these cases. It is somewhat puzzling that there are quite a few more out-of-business units in the telephone estimate group than in the telephone prompting group. They may reflect an inaccurate coding of Code 81, "Unable to locate: Employment ≤ 50 ." However, for response and estimation purposes, these codes are effectively the same.

These response rate measures are further analyzed by looking into the two important attributes of establishments: industry type and establishment size. Table 4 indicates that the results by industry type are similar to the overall results. It is interesting to note that the difference in response rates is considerably less in the manufacturing sector. This result may help substantiate results obtained in other tests indicating that the manufacturing sector has data which is easier to obtain and classify for the Hours at Work survey purposes than the non-manufacturing sector. On the other hand, response rates per 100 hours indicate that the telephone prompting approach produces a higher response rate in both industry groups than does the telephone estimate approach.

Table 5 also indicates that the results by establishment size are similar to the overall results. The telephone prompting approach produces response rates significantly lower than the telephone estimates across all size groups. It is noteworthy that the response difference is least among the medium size group (Employment is greater than 50 but less than 500) as we have seen the same pattern in a previous pilot test. Response rates of medium size groups have been traditionally better than those of small or large size groups, as they have the contact person assigned for survey responses, and this contact person rarely suffers from other competing tasks that would be demanded in small or large size establishments. Response rates per 100 hours also indicate that the telephone prompting approach produces a higher response rate across all size groups than the telephone estimate approach.

5. CONCLUSION AND FUTURE RESEARCH

The empirical evidence collected in the experiment indicates that the telephone prompting approach converts more non-responding units per hour than the telephone estimate technique. This result has not changed when controlling for industry type and size groups. With the prompting method, the data is more complete and accurate. However, there is a much higher incidence of refusals in the telephone prompting group. This was not a surprise as the increased respondent burden would lead us to expect this occurrence.

The experiment was however constrained by limited resources which have resulted in incomplete findings. Therefore, a more complete test should be conducted in the future which would allocate fairly equal resources across the two groups from the beginning to the end. It is also suggested that several interviewers be assigned to each follow-up group. Each interviewer should be assigned an equal number from each group, in order to control for interviewer difference.

Studies of follow-ups in establishment surveys are certainly a critical element to enhance our understanding of nonresponse and response that would be influenced by micro-variables, macro-variables, and meso-level variables. Micro-variables include survey

characteristics and attributes of establishment survey informants. Macro-variables encompass societal norms influencing the informant's perception of trust in the survey sponsor and confidentiality of data provided. Meso-level variables are those organizational characteristics moderating the informant's interaction with the interviewer and survey instrument. Nonresponse reduction, a conventional goal of follow-ups, should be weighed with the benefits that a follow-up technique would bring with regard to data quality and cost-effectiveness.

Note: The assistance of Kenneth W. Robertson in the data analysis is appreciated. The opinions expressed in this paper are those of the author and do not necessarily represent those of the US Bureau of Labor Statistics.

BIBLIOGRAPHY

- Chun, Y., and Robertson, K. (1996) "Reducing Nonresponse in Business Surveys," Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Dillman, D.A. (1978) *Mail and Telephone Surveys: the Total Design Method*, New York: John Wiley & Sons.
- Groves, R.M. (1989) *Survey Errors and Survey Costs*, New York: John Wiley & Sons.
- House, J.S., Gerber, W., McMichael, A.J. (1977) Increasing Mail questionnaire Response: A Controlled Replication and Extension, *Public Opinion Quarterly*, vol. 41, pp 95-99.
- Lyberg, L. and Dean, P. (1992) "Methods for Reducing Nonresponse Rates: A Review, paper presented at the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg.
- Rosen, R.J. and Clayton, R.L. (1992) "An Operational Test of FAX for Data Collection," Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Yammarino, F., Skinner, S., and Childers, T. (1991) "Understanding Mail Survey Response Behavior: A Meta-Analysis," *Public Opinion Quarterly*, Vol. 55, pp. 613-639.

Table 1. Hours Based Response Rates

Group	Usable Units	Hours Spent	Average Hours per Unit	Usable Units per 100 Hours	Average Response Rate per 100 Hours	Diff.
Prompt	120	380	3.17	31.6	6.1%	+1.3%
Estimate	218	875	4.01	24.9	4.8%	

* An asterisk indicates that the difference is significant at the alpha = 0.05 level.

Table 2. Response Rates

Group	Usable Units	Unusable Units	Not in Scope	Total units	Response Rate	Diff.	Hours Spent
Prompt	120	397	97	614	23.2%	-18.7%*	380
Estimate	218	302	94	614	41.9%		875

* An asterisk indicates that the difference is significant at the alpha = 0.05 level.

Table 3. Response Results

Response Code		Prompt	Estimate
0	Nonresponse	(2%) 13	(32%) 197
10	Pending	(30%) 183	(11%) 66
20	Refusal	(24%) 149	(5%) 33
40	Out of Scope	21	12
50	Out of Business	19	41
60	Data Unusable	17	0
70	Duplicate Unit	9	15
81	Unable to Locate: Employment <= 50	(6%) 35	(1%) 6
82	Unable to Locate: Employment > 50	(8%) 48	(4%) 26
90	Usable Unit	(20%) 120	(36%) 218
	Total	(100%) 614	(100%) 614
	Hours Spent	380	875

Table 4. Response Rates by Industry Level

Sector / Group	Usable Units	Hours Spent	Response Rate per 100hours	Response Rate	Diff.
Manufacturing					
Prompt	65	206	12.0%	26.2%	-12.6%*
Estimate	94	377	9.4%	38.9%	
NonManufacturing					
Prompt	55	174	12.4%	20.1%	-25.0%*
Estimate	124	498	9.8%	45.1%	

* An asterisk indicates that the difference is significant at the alpha = 0.05 level.

Table 5. Response Rates by Establishment Size

HW-Size / Group	Usable Units	Hours Spent	Response Rate per 100hours	Response Rate	Diff.
Emp<50					
Prompt	46	146	17.9%	26.1%	-22.0%*
Estimate	78	313	15.4%	48.1%	
50<=Emp<500					
Prompt	55	174	15.4%	26.8%	-12.8%*
Estimate	90	361	11.0%	39.6%	
500<=Emp<2500					
Prompt	16	51	3.2%	16.0%	-24.0%*
Estimate	40	161	2.5%	40.0%	
Emp =>2500					
Prompt	3	10	8.7%	8.3%	-23.9%*
Estimate	10	40	8.1%	32.3%	

* An asterisk indicates that the difference is significant at the alpha = 0.05 level.