Experiments to improve response rates among likely refusers on longitudinal surveys: Does assigning better interviewers and paying interviewer incentives work?

Paper prepared by Lisa Calderwood (Centre for Longitudinal Studies, Institute of Education), Hannah Carpenter (TNS-BMRB) and Andrew Cleary (Ipsos-MORI) for the 24th International Workshop on Household Survey Nonresponse, London, 4th-6th September 2013.

Introduction

One of the main challenges for longitudinal surveys is re-engaging respondents who have not participated in recent wave(s). However, one of the key advantages of longitudinal surveys is that information from previous waves can be used to identify sample members who are the least likely to take part in order to better target fieldwork interventions designed to maximize response for this group. This paper reports the results from two experiments designed to maximize response from the sample members classified as most likely to refuse on the most recent waves of the 1970 British Cohort Study (BCS70) and the UK Millennium Cohort Study (MCS). Both are large-scale longitudinal surveys. On BCS70 the fieldwork intervention involved assigning 'better' interviewers to the most difficult cases, identified using para-data from prior waves. On MCS the fieldwork intervention involved paying incentives to interviewers to convert the most difficult cases, identified using response propensity modeling. The main finding from these experiments was that these interventions did not lead to an increase in the response rates for these difficult cases, and hence they did not lead to a higher response rate overall. In the light of these findings, the paper will reflect on the nature of these interventions, the utility of targeting non-response interventions on difficult to co-operate groups, and the pros and cons of different approaches to targeting.

Experiment 1: Are 'better' interviewers more successful at engaging reluctant respondents?

Motivation

It is well known that some interviewers achieve better response rates than others. Groves and Couper (1998) argue that the causal mechanism explaining interviewer effects on nonresponse is most likely to be the behaviour of the interviewer during the interaction with the householder and the expectations of the interviewer about the likelihood of gaining co-operation. In the context of longitudinal surveys, the role of interviewer continuity in response maximization has been studied, with the majority of the research indicating that continuity is beneficial. However, much of this research does not adequately control for selection effects and the non-random allocation of interviewers to assignments. Moreover, in relation to re-engaging prior wave non-responders, recent evidence shows that using different interviewers is more successful than sending back the same interviewer (Watson and Wooden, 2013).

For most surveys, the allocation of sample members and assignments is based primarily on interviewer availability and geography, rather than interviewer ability. The skills of 'better' interviewers are often drawn upon at the re-issue stage for refusal conversion efforts among sample members who refuse the initial survey request. On cross-sectional surveys this is, in part, because it is difficult to know in advance which cases will be the most reluctant. However, on longitudinal surveys it is possible to use information from prior waves to identify sample members who may be more reluctant to respond, and then to allocate better interviewers to this group from the outset. This approach has the potential to save costs and boost response rates. All other things being equal, achieving a higher response rate from the

initial issue should reduce the proportion of cases requiring expensive refusal conversion attempts and save costs.

Data and methods

The aim of the experiment was to identify sample members who are most likely to refuse and then to allocate some of these cases to better interviewers and some to other interviewers in order to compare the response rate achieved by different types of interviewers among these cases. We hypothesized that better interviewers would achieve a higher response rate than other interviewers for these cases.

The experiment was conducted on the Age 42 survey of the 1970 British Cohort Study (BCS70) which took place in 2012-3. BCS70 is a large-scale longitudinal study following around 17,000 people who were all born in Britain in a particular week in 1970. The sample members have been approached on eight prior occasions: birth, 5, 10, 16, 26, 30, 34 and 38. The fieldwork as carried out by TNS-BMRB, one of the major survey research agencies in the UK.

Prior response patterns were used to identify the group considered most likely to refuse. Likely refusers were sample members who a) refused to take part in the last wave at age 38 in 2008 b) refused to take part in the age 34 wave in 2004, and also did not take part (although did not necessarily refuse) in 2008 c) refused to take part in the age 30 wave in 2000 and also did not take part (although did not necessarily refuse) in 2004 and 2008. It total 1,009 cases were identified in this way, which equated to 8.7% of the sample issued for the Age 42 Survey (n=11,654).

Better interviewers were identified using TNS-BMRB's interviewer performance index. This is an indicator based on interviewer response rates achieved on all assignments (across all projects) an interviewer has worked on in the last quarter. Interviewers were classed as "high performing" if they had an index score of **0.95** or higher. This means that on average they had achieved within 5% or lower of the target response rate on all of their assignments. In total, 274 interviewers worked on the age 42 survey and 70 of them (26%) were identified as high performers. Compared with other interviewers, high performing interviewers were more likely to be female, aged between 50-70 years and with between 2-10 years service.

The allocation of cases to interviewers was done geographically. Just under one-third (n=281) of likely refusal cases were allocated to better interviewers with the remainder (n=611) issued to other interviewers. (Note: total does not sum to 1,009 as cases withdrawn prior to fieldwork and re-issued were not included). Although the allocation of likely refusers to interviewers was not random, this led to a 'natural experiment'. Robustness checks (not included here) showed that there were no significant differences between likely refusers allocated to better interviewers and those allocated to other interviewers, on a range of relevant characteristics.

Results

Table 1 shows that response rates (and refusal rates) for the likely refusal cases were no different for high performing interviewers compared with other interviewers. Both types of interviewers were equally good at achieving an interview with the likely refusers (30% for high performing interviewers and 29% for other interviewers), with no significant differences in the other categories of non-response, including the refusal rate.

Table 1: Response rate for likely refusers by interviewer type

Outcome	High performers	Other interviewers	p-value
Interview achieved	**30%**	**29%**	0.76
Refusal	**51%**	**50%**	0.78
No contact	2%	3%	0.39
Untraced mover	11%	13%	0.40
No longer eligible (died/moved abroad)	3%	2%	0.36
Other unproductive	3%	4%	0.44
Base	*281*	*611*	

For other cases (not classified as likely refusals), higher performing interviewers did achieve slightly higher response rates (87% compared with 84%) and slightly lower refusal rates (7% compared with 9%) than other interviewers. There were no significant compositional differences between other cases assigned to the different types of interviewers. This suggests that the method of identifying better interviewers was robust.

These results also show that using paradata from previous waves on participation history is an effective way to identify likely refusers, with around 50% of these cases refusing to take part, compared with around 9% overall.

Discussion and implications

These results demonstrate clearly that better interviewers were no more effective at securing higher response rates from reluctant respondents than other interviewers, on a major longitudinal study in the UK.

It is not clear why better interviewers had no effect of response rates among reluctant respondents, particularly as they did achieve higher response rates overall. It may be that different skills are required to persuade reluctant respondents than securing co-operation generally, and that identifying better interviewers based on their ability to do this on longitudinal surveys, or on their re-issue performance, might have been more effective than the general performance indicator used. However, it may also be that the decision to re-engage in a longitudinal survey may be less susceptible to influence by the interviewer and driven primarily by respondents themselves, including their reasons for prior wave non-participation.

Experiment 2: Can interviewer incentives boost co-operation rates among reluctant respondents?

Motivation

There is lots of evidence that using respondent incentives can increase response rates and, in the context of a longitudinal survey, reduce attrition (Laurie and Lynn, 2009). However, there is relatively little evidence on the effect of interviewer incentives on survey response. In general, in the UK, interviewers are paid by results, rather than by hours worked. On most surveys, there are fixed payment amounts associated with achieving an interview, usually supplemented by payments for travel and the effort required to make contact and secure co-operation. However, for any given survey, the amount of effort required to secure an interview is variable, depending primarily on the contactability and willingness to participate of the sample member. Within this context, interviewers are arguably motivated to focus their efforts on achieving interviews with the easier cases. It is therefore hypothesised that paying interviewers additional incentives for difficult cases will increase their motivation to achieve

an interview, and lead them to make additional efforts for these cases. The implies that these extra efforts will lead to a higher co-operation rates overall than would otherwise have been achieved. There is very little empirical evidence about the effectiveness of interviewer incentives. Peytchev et al (2010) found that increased interviewer incentives were not successful at boosting response rates. As with respondent incentives, there are also equity concerns relating to the use of interviewer incentives, particularly in an experimental context involving differential payments to interviewers.

As noted earlier, for cross-sectional surveys, it is difficult to know in advance which cases are likely to require additional effort. However, on longitudinal surveys information from prior waves can be used to identify cases which are most likely to be difficult in advance, and fieldwork interventions, such as interviewer incentives, can be targeted on these cases.

Data and Methods

The aim of this experiment was to evaluate whether offering additional incentive payments to interviewers to secure interviews with reluctant respondents leads to higher co-operation rates for these cases than simply asking interviewers to make extra efforts to secure an interview with these cases. It was hypothesised that interviewer incentives would lead to higher co-operation rates among reluctant respondents.

The experiment was conducted on the Age 11 survey of the UK Millennium Cohort Study (MCS) which took place in 2012. MCS is a large-scale longitudinal study following over 19,000 families in the UK born in 2000/1. There have been four prior waves of the study: 9 months, 3, 5 and 7. The fieldwork as carried out by Ipsos MORI, one of the major survey research agencies in the UK.

Separate response propensity models, using both survey data and para-data from prior waves, were used to identify sample members who were least likely to be contacted and least likely to be to co-operate. The 25% of cases with the lowest co-operation propensity (after the bottom 10% in the contact propensity distribution had been removed) were selected as the target sample for the interviewer incentive experiment. These cases were allocated to interviewer assignments on a geographical basis, and then these assignments were randomly allocated to one of two phases of fieldwork which formed the experimental groups. These likely refusals were identified to interviewers as ‘target cases’. In the first phase of fieldwork (February 2012 start), which formed the control group for the experiment, interviewers were told that target cases were of ‘special value to the study’ and directed to make extra efforts to interview them. In the second phase of fieldwork (April 2012 start), which formed the treatment group, interviewers were given a £10 bonus for achieved interviews with these target cases. They were told that the bonus was ‘in recognition of the potential extra work required’. Robustness checks (not included here) show that the phase 1 and phase 2 samples were balanced on other relevant characteristics. The analysis was restricted to first issue cases only and to cases worked by interviewers who worked on both waves. The additional cost of the interviewer incentive was funded through a reduction in the budget available for re-issues.

Results

Table 2 shows co-operation rate for the target cases was no different for cases in the control group (no interviewer incentive) and the treatment group (£10 interviewer incentive). We also looked at overall response rate (not shown here), and found no difference between the

experimental groups. Table 2 also shows that the co-operation rates for non-target cases and for all cases also did not differ between the fieldwork phases.

Table 2: Co-operation rate for target cases by interviewer phase

	Phase 1 (control, no incentive)	**Phase 2 (treatment, incentive)**
Target cases	**72.2%**	**71.1%**
N	**1,106**	**1,174**
Non-target cases	90.4%	88.9%
N	3,936	4,136
All cases	86.5%	84.9%
N	5,042	5,310

We also examined paradata on interviewer calls to try to evaluate whether the introduction of the incentive lead to interviewers making more effort. We did not find any evidence of this. Finally, we compared the actual co-operation rate at wave 5 (age 11) by the predicted co-operation probability in deciles. This showed that the response propensities were a good predictor of co-operation behaviour at wave 5.

Discussion and implications

These results show that interviewer incentives were not effective at securing higher co-operation rates from reluctant respondents on a major longitudinal study in the UK. There was also no evidence that interviewers made greater efforts for these cases.

It is possible that the level of incentive was too low to make a difference. There are also other possible confounding factors which may have had an impact on the results, in particular other bonus payments for interviewers introduced during fieldwork and extensions to fieldwork meaning there was a greater degree of overlap between phases than was optimal (risking contamination between the experimental groups) and that all cases, though particularly phase 1 cases, had longer time in the field.

However, the main finding that the interviewer incentive did not make a difference indicates that within the context of a high-quality survey like MCS, with thorough fieldwork procedures and robust monitoring, the scope for interviewers making additional efforts to secure co-operation may be limited. It may also be that the decision to re-engage in a longitudinal survey may be driven primarily by respondents themselves, including their reasons for prior wave non-participation.

Conclusions and issues for discussion

Overall, these experiments show that it is possible to conduct robust evaluations of fieldwork strategies during the main stage fieldwork of large-scale surveys. They also show that it is possible to identify reluctant respondents using prior wave data on longitudinal studies and to target fieldwork interventions on them. The overall finding that neither better interviewers nor interviewer incentives had an impact on co-operation and response rates for reluctant respondents implies that alternative fieldwork strategies may be required to improve participation rates among these groups.

Issues for discussion:

- Is using response propensity modelling to identify potentially difficult cases worthwhile? Why not use simpler methods such as fieldwork outcomes from prior waves?

- What is the most appropriate cut-off for the identification of difficult cases in order to target interventions? And what proportion of all cases should be allocated to these interventions?
- Is it sufficiently robust to utilise 'natural' experiments in fieldwork organisation and allocations to evaluate different fieldwork strategies? How do we adequately account for 'real-world' confounders?
- How much potential is there to make a difference to response rates on high quality surveys? Aren't we already doing everything we can?
- What factors influence the decision to continue to participate in a longitudinal survey, or re-engage following non-participation? How can these be influenced by survey design and interviewers?

References

Groves, R.M. and Couper, M.P. (1998). "Nonresponse in Household Interview Surveys". New York: Wiley.

Laurie, H. and Lynn, P. (2009), "The Use of Respondent Incentives on Longitudinal Surveys". In P.Lynn (ed.) 'Methodology of Longitudinal Surveys', Chicester: Wiley

Peytchev, A., Riley, S., Rosen, J., Murphy, J. and Lindblad, M. (2010). "Reduction of Nonresponse Bias through Case Prioritization". *Survey Research Methods*, 4 (1).

Watson, N. and Wooden, M. (2013). "Re-engaging with survey non-respondents: evidence from three household panels", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *in press*.