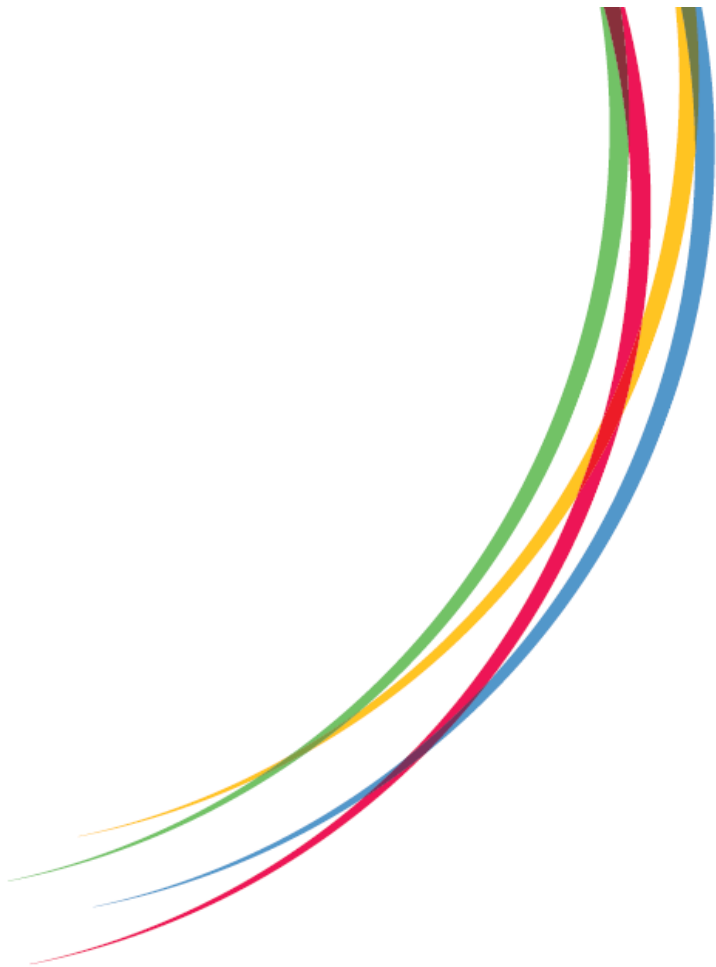




Understanding Society

Working Paper Series

No. 2013 – 02

March 2013

Targeted Response Inducement Strategies on Longitudinal Surveys

Peter Lynn

Institute for Social and Economic Research, University of Essex



Targeted Response Inducement Strategies on Longitudinal Surveys

Peter Lynn

Non-Technical Summary

It is very important for longitudinal surveys to achieve good response rates. This helps to ensure that the sample remains representative and also that it remains large enough to support longitudinal analysis. Survey researchers have long recognised that many aspects of the way a survey is carried out can affect response rates. Many experiments have been carried out to find ways of increasing response rates.

However, survey researchers tend to treat all members of a survey sample the same way. For example, if they believe that a well-designed letter, sent before an interviewer calls, will tend to increase response rates, they will send an identical letter to all sample members. There are good reasons to believe that different sample members will react differently to a survey feature such as a letter. Thus, the best strategy may not in fact be to send the same letter to everyone. It may be that different letters work better for different subgroups, or indeed that for some subgroups it is better not to send a letter at all.

An alternative to standard practice is therefore to treat sample subgroups differently, depending on what we know about best to achieve good participation in that subgroup. This is what we refer to as a “targeted response inducement strategy.” In this paper we outline what such a strategy consists of and we describe the steps involved in designing and carrying out such a strategy. The aim is to help people who carry out surveys to do so in a more effective way. We also review current practice in the use of such strategies and briefly present four examples.

Targeted Response Inducement Strategies on Longitudinal Surveys

Peter Lynn

Abstract

Targeted response inducement strategies involve varying features of survey design between sample subgroups in a way that is anticipated to achieve the best trade-off between costs and non-response. The features could include prenotification letters, incentives, between-wave mailings, website content, telephone calls, and so on. Such strategies are not much used on longitudinal surveys. We discuss why this might be the case and we describe the steps to identifying and implementing an effective strategy. We also present four examples.

Key words: balanced response, between-wave mailings, non-response, prenotification letters, respondent incentives, response rate

JEL classifications: C81, C83

Author contact details: plynn@essex.ac.uk

Acknowledgements: The work reported in this paper is supported by the UK Longitudinal Studies Centre at ISER, a UK Economic and Social Research Council (ESRC) resource investment.

Targeted Response Inducement Strategies on Longitudinal Surveys

Peter Lynn

1. Introduction

Two decades ago, survey researchers suggested that standardised interviewer introductions and requests for survey participation were not the most effective way to obtain the co-operation of sample members. Instead, they argued, interviewers should “tailor” their introductions to the circumstances and concerns of each sample member (Groves, Cialdini & Couper 1992, Morton-Williams 1993). The idea of interviewer tailoring quickly gained wide acceptance and is now a routine aspect of interviewer training and survey practice in all major survey organisations. However, the idea of adapting the nature of the survey protocol to the circumstances of sample members in order to improve response rates did not spread to other elements of the survey process. Instead, most other aspects of survey design and implementation typically remain standardised across all sample members. This is perhaps surprising, considering that a sizeable methodological literature has demonstrated that the effects of various design features on response rates are often uneven across sample subgroups. In other words, some strategies are more effective for some types of people than for others.

More recently, researchers have begun to explore the idea of treating sample subgroups differently in order to achieve an appropriate trade-off between survey costs, on the one hand, and some error or quality indicator(s) on the other hand. The indicators typically relate either to measurement or to non-response. Interest has focused on the idea of starting data collection in a standardised way but then changing it in different ways for different sample members as fieldwork progresses. The changes can either be targeted at one or more subgroups, as in the responsive designs of Groves and Heeringa (2006), or tailored to individual sample members, as in the adaptive designs of Wagner (2008). A broad distinction can be made between targeted and tailored strategies. Targeted strategies involve treating each of a limited number of sample subgroups in different ways, where the subgroups are defined by relevant common characteristics. Tailored strategies involve treating each individual sample member differently, depending on that person’s

characteristics and behavior (Couper and Wagner 2011). At the margins, the distinction is difficult to pin down, but this paper focuses on targeting rather than tailoring.

The distinction between responsive and adaptive designs is not clear and Couper and Wagner (2011) have argued that the two lie at different points on a continuum of possible design variants. Schouten et al (2013) additionally include designs in which features are targeted at subgroups from the start – rather than only as fieldwork progresses – under the rubric of adaptive designs. It is designs of this kind with which we are concerned in this paper – and specifically those in which the relevant quality indicators are either response rate or response balance.

We review both theory and practice regarding targeted strategies to improve response rate or response balance. We define targeted strategies broadly and in section 4 below we outline the different ways in which strategies can be targeted and identify three main categories of targeted strategy.

2. Response Inducement

Survey researchers have long been concerned with how best to obtain good response rates and have developed many elements of the survey process in ways intended to increase response rates. Such elements include, amongst others, the design of survey materials, the substantive content of survey communications, the nature and intensity of field effort and the provision of direct incentives such as money, gifts, or lottery tickets. At every step in the survey design and implementation process, survey researchers consider carefully the likely implications of each possible decision for survey non-response. Often these considerations conflict with other objectives such as cost-efficiency, timeliness and parsimony. For this reason, the efficacy of measures to reduce non-response must be well-established in order for researchers to be able to justify incorporating a measure that costs more or delays completion of fieldwork, for example.

Large research literatures address some of the survey design features that affect response rates. Features found to affect response rates include the form and value of monetary incentives offered to respondents (Church 1993, Singer et al 1999, Singer 2002, VanGeest et al 2007), the content and design of prenotification and invitation letters (de Leeuw et al 2007, Dillman 2007, Hembroff et al 2005, Kaplowitz et al 2012), mode of prenotification (Bosnjak et al 2008, Kaplowitz et al 2004) interviewer calling patterns (Bennett & Steel 2000, Campanelli et al 1997, Weeks et al 1987), survey topic and sponsorship (Groves et

al 2004, Sheehan 2006, Van Kenhove et al 2002), respondent burden (Crawford et al 2001, Galesic & Bosnjak 2009) and interviewer characteristics (Beerten 1999, Durrant et al 2010, Jäckle et al 2013). Some of these features appear to be relevant for all survey modes, while others are specific to certain modes.

However, all of these findings relate to sample mean effects. On average, each of these features tends to affect response rates, so choice of the most effective option will tend to increase response rates, but this is not uniformly true for all sample subgroups. Indeed, such heterogeneity of effect may explain some of the inconsistencies in the methodological literature. If a particular feature improves response amongst older people, but not amongst younger people, for example, you would expect to observe a sample mean effect on a study of the total population, a stronger effect on a study of retired people, and no effect on a study of students. And there is plenty of evidence in the literature of heterogeneous effects on response propensity. For example, Kaplowitz et al (2012) found that the length of the invitation letter affected response rates to a web survey for staff and faculty of a university but not for students, while information about how long the survey would take to complete affected response rates for students, but not for faculty or staff. Similarly, Lynn (2012a) found that a monetary incentive improved response rates for older people but not for younger people. But rather than being viewed as a limitation of existing research, this heterogeneity of effects can be considered as an opportunity to achieve positive effects that are even stronger than observed sample mean effects. In principle, this can be achieved by targeting strategies to the sample subgroups where they are most effective.

Such targeting could have either of two objectives. One possible objective would be to maximize the survey response rate. An alternative would be to achieve balanced response across important covariates (for example, maximising an R-indicator in the manner of Schouten et al, 2009). In either case, this should be achieved in a cost-effective way. The two objectives can have different implications for the nature of the targeting, but in either case the basic strategy is the same.

In practice, targeted response inducement strategies are not often used. There are a number of reasons for this:

- There is a lack of information about how treatment effects vary over sample subgroups. Many articles reporting the effects of features designed to boost survey response report only the sample mean effects, with no information provided regarding the heterogeneity of the effects.

- Methodological studies define subgroups inconsistently. Where effects on subgroups are reported, the subgroups are defined differently in different studies, making it difficult to combine information to inform a strategy involving multiple design features.
- Often, subgroup membership is not known in advance. Many sampling frames contain little or no information that can be used to target response inducement strategies. This is particularly true for address-based sampling, where geographical location is typically the only characteristic known in advance. If the subgroup membership of a sample member is not known in advance, response inducement strategies cannot be targeted.
- Subgroups to be targeted may not correspond to subgroups for whom effect estimates are available. Even when a sampling frame is informative, the available variables may not be ones which any previous methodological study has investigated.
- The features available for manipulation are limited. The opportunities to influence the participation decision are typically limited to one or two brief communications. When these communications are written rather than interviewer-administered, and especially in the case of address-based sampling and similar situations in which the respondent is not yet identified, the communication may not be seen at all by some sample members.

Happily, some of these barriers to the use of targeted response inducement strategies are reduced in the case of longitudinal surveys. In the next section we discuss why this is the case and how we can take advantage of the opportunities provided by longitudinal surveys for the implementation of targeted strategies.

3. Longitudinal Surveys

A strength of longitudinal surveys is the availability, once the recruitment wave has been completed, of a wide range of pertinent measures for each sample member. These rich data provide an opportunity to identify a multitude of alternative subgroups at whom response inducement strategies can be targeted. This opportunity is enjoyed only by longitudinal surveys. Even the most informative of sampling frames such as population registers and administrative data bases typically include only a few variables, and not in as much detail as can be collected in a survey. Additionally, the richness of the information available in a longitudinal survey often means that measures can be constructed that correspond closely to the subgroup definitions used in earlier methodological studies, enabling direct projection of methodological findings to the survey sample.

A further difference between longitudinal surveys and cross-sectional surveys is that there are typically more features available for manipulation in the case of longitudinal surveys.

This too makes the context for targeted response inducement strategies more promising in the longitudinal case. With a cross-sectional telephone survey, manipulable features are limited to the prenotification letter and the interviewer introduction. With a cross-sectional face-to-face survey, there may also be materials designed for the interviewer to hand over or show to the respondent. And with a mail survey the design of the questionnaire itself can also be manipulated. All of those possibilities apply to the longitudinal case too, but additionally there may be various kinds of between-wave interventions, including post-wave thank you letters, between wave motivational mailings of findings or other information, keep-in touch mailings or phone calls, birthday cards, change-of-address cards, and so on. Additionally, longitudinal surveys are increasingly developing participant websites where sample members can log in, read information of various kinds and take part in various activities. The nature of these web-based interactions could also be targeted.

It could also be noted that there is perhaps a greater justification in the case of longitudinal surveys for the implementation of measures such as targeted inducement strategies that can improve overall response rates. This is simply because response rates are in some sense more important for longitudinal surveys than for other surveys. In addition to being a contributing factor to non-response bias (and a rather crude indicator of the likelihood of bias: Groves & Peytcheva 2008), as for any survey, the continuing value and utility of a longitudinal survey relies on the continued participation of a high proportion of the sample members. Unlike the cross-sectional case, a loss of sample numbers cannot be compensated by simply adding more sample, as the historical longitudinal data will be missing and cannot be fully recreated retrospectively. There is therefore a particular imperative for high wave-on-wave response rates for longitudinal surveys. Without them, the whole survey becomes unviable.

4. The Targeted Inducement Design

There are a number of ways in which inducements could be targeted. It may be useful to distinguish between three main categories:

- A design feature is applied only to one group. This could be done because the feature is believed to be effective only for that group, or because it is only appropriate for that group, or simply because that group is otherwise expected to have a lower response rate than other groups;

- A different design feature is offered to each group. Such a design may be chosen because it is believed that the optimum feature differs between groups, for example where there is a choice between two incompatible features such as two different types of incentive;
- A group-specific variant of the same design feature is offered to each group. An example would be a prenotification letter in which the wording is varied between groups.

In any of these three cases, the tasks facing a researcher who wishes to introduce such a design consist of the following components:

1. Defining the subgroups to be targeted;
2. Identifying the inducement strategy to be applied to each subgroup;
3. Allocating sample units to subgroups;
4. Implementing the design.

Note that we refer here specifically to targeting that takes place before a wave of data collection begins, based on prior characteristics. This is distinct from the practice of adapting design features in response to respondent behaviour during the course of a wave of data collection. For example, on the American Panel Study of Income Dynamics (PSID; McGonagle et al 2012), respondents who are interviewed on their cell phone and who complain about the cost to them of the cell phone minutes are offered an additional \$10 payment in compensation (McGonagle 2013). Although this additional payment goes to only a subset of sample members, we do not consider this a targeted design as the payment would be offered to *any* sample member who complains about the cost of receiving a call on their cell phone. Similarly, the Household, Income and Labour Dynamics in Australia (Watson & Wooden 2012) survey sends a letter to households who have not participated by the end of the first field work period within each wave (Watson 2013). There are different versions of the letter, depending on the reason for not yet participating. Again, however, we do not consider this a targeted design as any sample household would receive the same treatment if they had the same outcome at the end of the first field period. Many other examples of such reactive adaptation on the part of the survey organisation can be found. One can even think of the persuasion strategies used by interviewers and the choice of when to make the next call attempt to a sample member (whether that be determined by an interviewer or by a call scheduling algorithm) as examples of this.

We now outline in turn the requirements for each of the four components of a targeted inducement design.

Defining the subgroups

The subgroups must be ones for which distinctive and effective inducement strategies can be identified. In practice this means that the first two components of the design listed above cannot be carried out independently. Simultaneous consideration of the two issues is needed.

If the objective is not only to increase response rate but also to improve response balance, then each subgroup should be relatively homogeneous both in terms of survey response propensity and in terms of key survey variables. In other words, the subgroups should be ones that are likely to differ from one another in terms of response rate in the absence of any targeted design and there should also be variation between the subgroups in the means, distributions and associations of key variables. These requirements are somewhat analogous to the requirements for successful cell-weighting for non-response adjustment. This is not surprising if we consider that in both cases the objective is to reduce non-response bias. The difference is that targeted inducement designs aim to reduce bias at source at the data collection stage, whereas weighting adjustments aim to reduce bias post-hoc at the estimation stage.

However, there are also two important differences between the requirements for subgroup definition for targeted inducement designs and the requirements for effective weight-adjustment cells. The first is the need to be able to develop distinctive and effective inducement strategies, as mentioned above. Second, there is no particular need to avoid small subgroup sample sizes as the approach does not involve estimation for subgroups (though designing targeted materials for very small numbers of sample cases may be considered inefficient and researchers may prefer to avoid small groups for that reason).

Identifying targeted inducement strategies

The optimum strategy for each subgroup is the one that maximises the expected response rate, subject to budget constraints. However, a desire to achieve balanced response would suggest that resources for response inducement should be devoted disproportionately to those subgroups that have the lowest predicted response rates in the absence of a targeted strategy (Calinescu et al 2013). Thus, strategies with a higher unit cost of implementation can be considered for the lowest response propensity groups if this is likely to produce a larger response rate gain for those groups than for other groups. Methodological experimentation is needed in order to establish the relative effectiveness of possible response inducement strategies for different subgroups. Previously published

studies of design features that influence response propensity can therefore be used to inform the choice of targeted strategy, but only if those studies have included subgroup analysis. Alternatively, researchers may wish to carry out their own methodological tests. In many ways longitudinal surveys are ideal vehicles for this as an experiment mounted at one wave can inform the design at the next wave. For example, a particular design feature could be implemented for a random half of the sample units at one wave. Analysis of the outcomes at that wave should reveal the subgroups for which the feature was most effective. At the following wave, this feature could then be implemented only for those subgroups. Assuming that there is a resource cost associated with implementing the feature, this should be more cost-effective than simply implementing the feature for the whole sample.

In the absence of information about difference in performance of features across subgroups, a targeted design can still be implemented, either by drawing upon non-response theory or through selective use of design features that have been shown to have a mean effect. For example, if a costly design feature (such as a large respondent incentive, for example) is known or believed to have a positive effect on response for all subgroups, this could be applied just to subgroups expected to have low response propensity, even in the absence of any evidence that it would be any more effective for those subgroups than for others.

Allocating sample units to subgroups

The subgroup definitions must be applied to sample units based on information available either from the sampling frame or, more likely, from data collected at previous waves of the survey. This can be a very straightforward process if variables defining the subgroups are available for all sample units (no missing data) and if these measures are accurate (no measurement or processing errors). More often it may be the case that the available variables only provide approximations to the ideal definition of the subgroups and that there is item missing data to deal with. In this situation the distinction is important between design features that are subgroup-specific and those which are generally applicable but applied selectively. If the design features under consideration are all of the generally applicable type, then a degree of misclassification of sample units to subgroups may not be of great importance. But if subgroup-specific features are to be used, care should be taken to avoid upsetting or annoying sample members by providing information, materials or persuasive messages that are inappropriate. For example, if a prenotification letter includes messaging that is specific to sample members with dependent children, a person

without children may be upset at the implication that the survey organisation believes they have children. The targeted features could then damage co-operation propensity rather than enhance it.

Implementing the design

The additional complications of implementing a targeted design are unlikely to be substantial but should not be underestimated. Care must be taken to ensure that the correct treatments are applied to the correct sample members. The derived indicator of subgroup membership will determine the treatment applied and should therefore be subject to rigorous quality control checking. Similarly, the programming of treatments depending on this indicator – such as the generation of mailings or the appearance on screen of scripts for interviewers – needs to be carefully tested. Where interviewers are involved in delivering the treatment, appropriate briefing is required to ensure that interviewers are aware that the treatment will vary between sample members and of the importance of delivering the allocated treatment.

5. Current Practice and Methodological Research

McGonagle et al (2011) provide an example of a methodological experiment carried out on the PSID in which a heterogeneous effect on field outcomes is observed. Specifically, they show that two features of a between-wave mailing reduce the number of calls needed at the following wave, but only for sample members who had indicated at the previous wave that they were likely to move in the next couple of years¹. The two features in question were offering an unconditional rather than conditional incentive to reply to a contact information request and using a traditional rather than new design for the mailing. These results may appear to suggest a targeted design in which the unconditional incentive and traditional design are used only for sample members with a high likelihood of moving. However, for the rest of the sample these features are no less effective than the alternatives. In such a situation, the organisational simplicity of having a standard rather than targeted mailing may suggest this to be the most efficient option, particularly if the cost differential between the alternative treatments is modest. PSID has not adopted a

¹ The PSID between-wave interval is two years.

targeted strategy for between-wave mailings. There are few examples of targeted inducement designs being used in practice. We describe here four examples.

Example 1: Prioritising Contact on Wave 5 of the UK Millennium Cohort Study

At wave 5 of the UK Millennium Cohort Study (Hansen 2012), sample households that were predicted to have the highest propensity to be not contacted (including not located) were issued to field first in order to maximise the time available in field and therefore hopefully increase the contact rate. These households were identified by fitting models predicting contact propensity at wave 4 based on outcome history at previous waves of the survey, household and neighbourhood characteristics (Calderwood et al 2012): the top 10% in rank order of predicted propensity to be not contacted were identified for the targeted treatment, along with any households known to have moved since the previous wave.

Example 2: Targeted Between-Wave Mailings on *Understanding Society*: the UK Household Longitudinal Study

On *Understanding Society*, each sample member receives a mailing approximately six months after each annual interview. This mailing includes a ‘thank you’ letter, a leaflet highlighting some survey findings, and a change-of-address card. For the post-wave 2 mailing, six versions of the leaflet were utilised: five versions each aimed at a target subgroup (young people, employed people, older people, ethnic minorities, and people who had previously been members of the British Household Panel Survey sample and therefore had a much longer association with the survey than others) and a general version sent to all other sample members. This approach was introduced following methodological experimentation which suggested that the targeted approach could improve response rates (Fumagalli et al 2013). This experimentation focussed on two subgroups, younger people and work-busy people, and the main findings are summarised in Table 1. It can be seen that the targeted version of the leaflet seems to encourage young people to complete a full face-to-face interview rather than a shorter telephone interview, and encourages work-busy people to at least complete the shorter telephone interview rather than not participate at all.

Table 1: Effect of a targeted leaflet on response for two subgroups

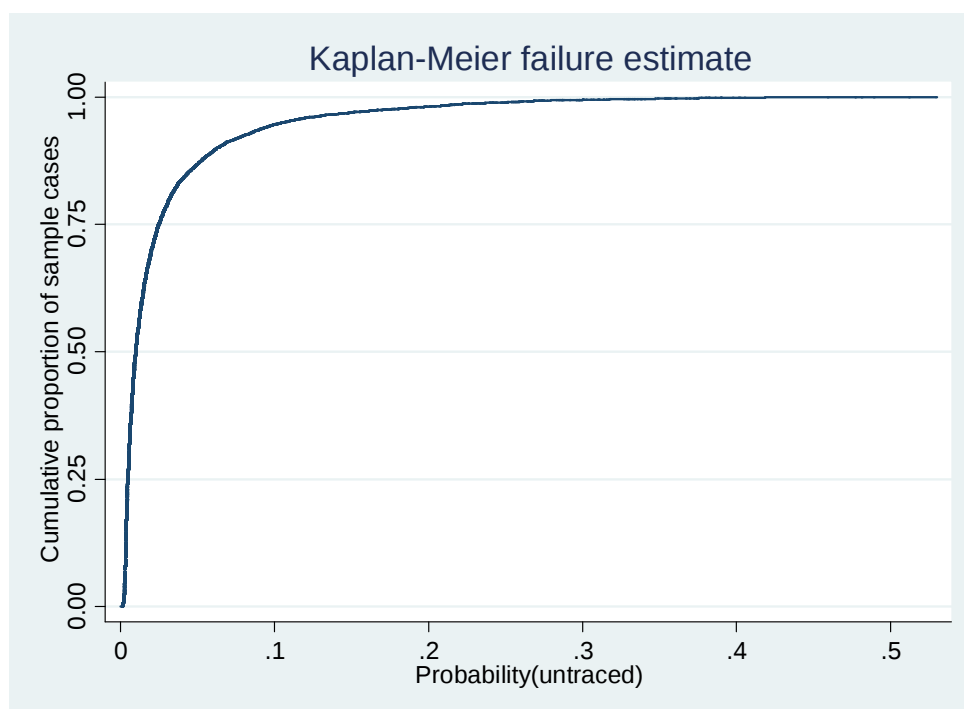
		Young			Work-busy		
		Standard	Targeted	<i>P</i>	Standard	Targeted	<i>P</i>
		%	%		%	%	
Full face-to-face interview		91.6	93.2	.106	90.1	90.3	.435
Short	telephone interview	2.6	0.8	.002	6.5	7.2	.250
Non-response		5.8	5.9	.465	3.5	2.5	.077
<i>n</i>		856	843		1,157	1,205	

Notes: The base consists of all sample members issued to field for BHPS wave 18. This includes all those enumerated at wave 17, aside from any known to have died subsequently, plus wave 17 non-respondents who had responded at wave 16. Short telephone interviews are offered as a last resort, only when all usual efforts to achieve a full face-to-face interview have been made. Young people are those who were aged 16–24 years at the start of wave 18 field work. Work-busy people are those who reported working more than 40 hours per week, or commuting for more than 10 hours per week in addition to working full time, or being self-employed. For more details of the experimental design, see Fumagalli et al (2013).

Example 3: Extra Keeping-in-Touch Mailings Between Waves 4 and 5 of *Understanding Society*: the UK Household Longitudinal Study

An additional between-wave mailing, asking sample members to notify any recent changes of addresses or imminent moves, was sent to those predicted to be at highest risk of moving and remaining untraced at wave 5. Predictions were based on wave 4 responses, using a model of wave 2 outcome based on wave 1 responses (Lynn 2012b). High risk was defined as a predicted probability of greater than 7.5%, a criterion that applied to 8.1% of sample members (so 91.9% of sample members did not receive this extra mailing). Figure 1 shows that this cut-off point represents a turning point in the distribution of predicted probabilities, so one should expect the targeted group to be somewhat distinctive. During the period of approximately 11 months between the wave 4 interview and the wave 5 prenotification letter, the majority of sample members received just one mailing (as described in example 2 above), while the targeted group received two. Thus, the mean interval between contacts was reduced from 5.5 months to 3.7 months for this group, increasing the chances of a change of address being identified before it was too late.

Figure 1: Predicted probabilities of remaining untraced at UKHLS wave 5



Example 4: Targeted Prenotification Letters on the *Understanding Society* Innovation Panel

An experiment is being carried out at wave 6 of the *Understanding Society* Innovation Panel (Uhrig 2011), in which targeted versions of the prenotification letter are being sent to a random half of the sample. The five target groups are younger people, people with dependent children, work-busy people, people of pensionable age and people living in London and south-east England. Each version of the letter mentions reasons why the survey is particularly important for the respective subgroup. A sixth version of the letter is sent to residual sample members who do not belong to any of the five target groups. With this design – as with example 2 – it is necessary to define the subgroups based on a hierarchy of criteria. For example, sample members who have dependent children and are also work-busy are placed in the work-busy subgroup rather than the dependent children subgroup. This is based on the observation that re-interview rates are lower overall for work-busy people than for people with children, so arguments for participation by work-busy people may be more important for this group.

6. Looking Forward

Targeted response inducement strategies seem not to be common on longitudinal surveys, though they have been discussed in international forums, including the International Panel Survey Methods Workshop in both 2010² and 2012³. There are several possible reasons for this. One is that they add extra complexity to the administration of surveys that are invariably already complex. This is undoubtedly true, though for many of the types of design features that are likely to be manipulated with such strategies, the extra work involved may be fairly minimal. Another concern could be the cost of implementation, but if designed well, a targeted strategy could significantly improve cost-effectiveness. Even a simple mailing to sample members can be quite costly for surveys with large sample sizes. But if the mailing could be targeted just to the subgroup for whom it is effective, or where the effect is most needed, considerable efficiency savings could accrue. More expensive design features, such as large incentives or gifts, might be ruled out for the whole sample, but could become affordable (and remain effective) if well targeted.

Ethical considerations may be a barrier to some kinds of targeted strategies. Review boards and ethics committees may not take kindly to the idea of different sample members receiving different kinds of rewards for participating in a study. This could affect the use of targeted respondent incentives, though it is hard to imagine the argument being applied to letters, leaflets or booklets. In any cases, researchers should certainly consider carefully the ethical implications of any proposed targeted strategy.

But in reality the main barrier to the introduction of targeted response inducement strategies may simply be a lack of experience of them. As more experience is gained, and more evidence is gathered regarding the types of designs that can be effective, and the circumstances in which they are more likely to be effective, the popularity of such designs is likely to grow. For this reason, the experimental introduction of such designs – as in example 4 above and as in Fumagalli et al (2013) – is to be encouraged. Experiments of this kind allow the research community to learn a lot about how to improve the trade off between survey costs, on the one hand, and response rates and non-response bias, on

² <http://www1.mzes.uni-mannheim.de/konf/psmw2010/>

³ http://melbourneinstitute.com/miaesr/events/workshops/workshop_panel_2012.html

the other hand. In the future, targeted strategies may become the norm for longitudinal surveys.

References

- Beerten, R. (1999). The effect of interviewer and area characteristics on survey response rates: an exploratory analysis. *Survey Methodology Bulletin* 45, 7-15.
- Bennett, D.J. & Steel, D. (2000). An evaluation of a large-scale CATI household survey using random digit dialling. *Australian and New Zealand Journal of Statistics*, 42(3), 255-270.
- Bosnjak, M., Neubarth, W., Couper, M. P., Bandilla, W., & Kaczmirek, L. (2008). Prenotification in web-based access panel surveys: The influence of mobile text messaging versus E-mail on response rates and sample composition. *Social Science Computer Review*, 26, 213-223.
- Calderwood, L., Cleary, A., Flore, G. & Wiggins, R.D. (2012) Using response propensity models to inform fieldwork practice on the 5th wave of the Millenium Cohort Study. Paper presented at the *International Panel Survey Methods Workshop*, Melbourne, Australia, July.
- Calinescu, M., Bhulai, S. & Schouten, B. (2013) Optimal resource allocation in survey designs. *European Journal of Operational Research*, 226, 115-121.
- Campanelli, P., Sturgis, P. & Purdon, S. (1997). *Can you hear me knocking? An investigation into the impact of interviewers on survey response rates*. The Survey Methods Centre at SCPR, London, GB. Available at <http://eprints.soton.ac.uk/80198/>.
- Church, A.H. (1993) Estimating the effect of incentives on mail survey response rates: a meta-analysis. *Public Opinion Quarterly*, 57(1), 62-79.
- Couper, M.P. & Wagner, J. (2011) Using paradata and responsive design to manage survey nonresponse. *Proceedings of the 58th World Statistical Congress, Dublin*, 542-548. Available at <http://2011.isiproceedings.org/papers/450080.pdf>.
- Crawford, S.D., Couper, M.P. & Lamias, M.J. (2001). Web surveys: perceptions of burden. *Social Science Computer Review* 19(2), 146-162.

- De Leeuw, E. D., Callegaro, M., Hox, J. J., Korendijk, E., & Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys: A meta-analysis. *Public Opinion Quarterly*, 71, 1-31.
- Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method* (2nd ed.). New York, New York: John Wiley.
- Durrant, G.B., Groves, R. M., Staetsky, L., & Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly* 74, 1-36.
- Fumagalli, L., Laurie, H. & Lynn, P. (2013) Experiments with methods to reduce attrition in longitudinal surveys. *Journal of the Royal Statistical Society Series A (Statistics in Society)* 176(2), 499-519.
- Galesic, M. & Bosnjak, M. (2009) Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349-360.
- Groves, R.M. & Heeringa, S.G. (2006) Responsive design for household surveys: tools for actively controlling survey nonresponse and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169 (3): 439-457.
- Groves, R.M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opinion Quarterly*, 72(2), 167-189.
- Groves, R.M., Cialdini, R.B. & Couper, M.P. (1992) Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475-495.
- Groves, R.M., Presser, S. & Dipko, S. (2004) The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, 68(1), 2-31.
- Hansen, K. (2012) (ed.) *Millennium Cohort Study First, Second, Third and Fourth Surveys: A Guide to the Datasets (Seventh Edition)*. London, Centre for Longitudinal Studies.
- Hembroff, L. A., Rusz, D., Rafferty, A., McGee, H., & Ehrlich, N. (2005). The cost-effectiveness of alternative advance mailings in a telephone survey. *Public Opinion Quarterly*, 69, 232-245.

- Jäckle, A., Lynn, P., Sinibaldi, J., & Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. *Survey Research Methods*, 7(1), 1-15.
- Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly*, 68, 94-101.
- Kaplowitz, M.D., Lupi, F., Couper, M.P., & Thorp, L. (2012). The effect of invitation design on web survey response rates. *Social Science Computer Review*, 30(3), 339-349.
- Laurie, H., & Lynn, P. (2009). The use of respondent incentives in longitudinal surveys. Chapter 12 in P. Lynn (ed.), *Methodology of Longitudinal Surveys*, Chichester: John Wiley.
- Lynn, P. (2012a). The propensity of older respondents to participate in a general purpose survey. *Understanding Society Working Paper 2012-03*, Colchester: University of Essex. At <https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2012-03>
- Lynn (2012b) Failing to locate panel sample members: minimising the risk. Paper presented at the *International Workshop on Household Survey Nonresponse*, Ottawa, Canada, September.
- McGonagle, K., Couper, M.P. & Schoeni, R.F. (2011) Keeping Track of Panel Members: An Experimental Test of a Between-Wave Contact Strategy. *Journal of Official Statistics* 27(2), 319-338.
- McGonagle, K., Schoeni, R.F., Sastry, N. & Freedman, V.A. (2012) The Panel Study of Income Dynamics: overview, recent innovations, and potential for lifecourse research. *Longitudinal and Life Course Studies* 3(2), 268-284.
- McGonagle, K. (2013) Personal communication. 17-02-2013.
- Morton-Williams, J. (1993). *Interviewer Approaches*. Aldershot: Dartmouth.
- Schouten, B., Calinescu, M., & Luiten, A. (2013, forthcoming). Optimizing quality of response through adaptive survey designs. *Survey Methodology*.
- Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology* 35(1), 101-113.

- Sheehan, K.B. (2006). Email survey response rates: a review. *Journal of Computer-Mediated Communication* 6(2).
- Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys, 163-177 in *Survey Nonresponse* (ed.s R.M.Groves, D.A.Dillman, J.L.Eltinge and R.J.A.Litle), New York: Wiley
- Singer, E., VanHoewyk, J., Gebler, N., Raghunathan, T. & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics* 15(2), 217-230.
- Uhrig, S. C. N. (2011) Using experiments to guide decision making in Understanding Society: Introducing the Innovation Panel, chapter 13 in S. L. McFall & C. Garrington (ed.s), *Understanding Society: Early Findings from the First Wave of the UK's Household Longitudinal Study*, Colchester: University of Essex. At: <http://research.understandingsociety.org.uk/findings/early-findings>
- VanGeest, J.B., Johnson, T.P. & Welch, V.L. (2007). Methodologies for improving response rates in surveys of physicians: a systematic review. *Evaluation and the Health Professions* 30(4), 303-321.
- Van Kenhove, P., Wijnen, K. & De Wulf, K. (2002). The influence of topic involvement on mail survey response behavior. *Psychology & Marketing* 19(3), 293-301.
- Wagner, J. (2008) Adaptive Survey Design to Reduce Nonresponse Bias, PhD thesis, University of Michigan, USA. Available at http://deepblue.lib.umich.edu/bitstream/handle/2027.42/60831/jameswag_1.pdf
- Watson, N. (2013) Personal communication. 19-02-2013.
- Watson, N. & Wooden, M. (2012). The HILDA survey: a case study in the design and development of a successful household panel study. *Longitudinal and Life Course Studies* 3(3), 369–381.
- Weeks, M.F., Kulka, R.A. & Pierson, S.A. (1987) Optimal call scheduling for a telephone survey. *Public Opinion Quarterly*, 51(4), 540-549.