Targeted letters: Effects on sample composition and item non-response

Annamaria Bianchi* and Silvia Biffignandi
Department of Management, Economics and Quantitative methods, University of Bergamo, Bergamo, Italy

Abstract. General population surveys as well as business surveys are affected by low response rates. There is a need to investigate possible ways to increase participation in surveys and to motivate interviewee in answering questionnaires. Even in official statistics the problem is relevant since lack of motivation undermines not only participation to the survey, but it is related to the timeliness in participation, the quality of the answers and hence the quality of the measurements of socio-economic phenomena. A possible approach to cope with this problem, which has received attention in recent years, is targeted design. In targeted designs, several survey features may be the object of interventions (timing of contact, number of contact attempts, mode of contact, interviewers' assignment, communication features, etc.). In our paper, we consider the invitation letter, which is one possible targeted feature. This is related to motivation to respond to surveys and to respond accurately, completely, and timely. We analyze a case study in the UK Household Longitudinal Study Innovation Panel. At wave 6 of the panel, a randomized experiment was carried out to study the effects of targeted initial letters [1]. We investigate effects of targeted initial letters on sample composition, representativity, and item non-response.

Keywords: Data quality, initial letters, adaptive designs, response, contact, subgroups

1. Introduction

To cope with the problems caused by non-response, recently, researchers have explored the idea of treating sample subgroups differently and intervening in the data collection process in order to achieve a "better" response set. Different terms are used to identify these methods, such as targeted design, responsive design, adaptive design, tailored design [2–4]. They all refer to slightly different versions of the method. In general, these methods are a response to an increasingly difficult survey climate characterized by declining budgets, rising costs and falling response rates [5]. Using these methods, the researcher tries to achieve goals such as variance minimization, non-response bias minimization, response rate maximization or even measurement error reduction. See [6] for an overview of the literature on the variety of approaches, on possible applications and on a critical discussion of some experimental studies.

Targeting survey features to sample units is currently a topic of research at National Statistical Institutes (NSIs). Experiments have been performed by NSIs mainly in the cross-sectional context (e.g. [4,7,8]). In general, the method proved to be useful. However, it is not standard practice yet. [9] provides theoretical and empirical evidence that a more balanced response achieved by the implementation of adaptive survey designs coincides with less non-response bias, even after adjustment.

In general targeted designs are focused on the reduction of biases while several other aspects have received little attention in the literature. Targeted design could have side effects on quality indicators other than those considered in the implementation of the design. This is related to the complexity of the concept 'data quality'. Many characteristics of the collection process affect data quality and, as a consequence, various quality measures should be used and the interaction among indicators should be evaluated in a Total Survey Error (TSE) perspective. This is even more important nowadays that changes in data collection tools and processes are currently being undertaken. IT diffusion is changing survey tools and the context where surveys are administered, the modes used in data collection, and

---
*Corresponding author: Annamaria Bianchi, Department of Management, Economics and Quantitative methods, University of Bergamo, Bergamo, Italy. E-mail: annamaria.bianchi@unibg.it.

the interaction among data quality aspects. As a consequence, innovative data collection ways also allow for continuous improvement and for monitoring better quality achievement.

In this paper, the focus is on manipulations to motivate interviewee in answering questionnaires. This can be achieved by interventions on initial letters. In interviewer administered surveys, the letters take the form of prenotification letters: sample members are sent letters notifying them that an interviewer will be calling. In mail and web surveys, they are invitation letters, inviting sample members to complete the survey. In both cases, initial letters generally describe the purpose of the survey, its length, and provide information on incentives, the legitimacy of the study and privacy issues. The purpose of these letters is mainly to motivate cooperation [10]. The influence can come from different sources: recognition of an authority [11], appeal to altruism [12], social validation [13], softening concerns regarding the use of the data [14], and interest in the survey topic [15]. The focus here is on interest in the topic of the survey. Sample members who are interested in the topic of the survey have a higher chance to respond than someone who is not interested [16,17]. Further, [18] finds that the effect of follow-up procedures is larger for those less interested in the survey topic. These effects may be related to the leverage-saliency theory [19], which postulates that features of the survey design have different impact on different kinds of people. Suitably-worded targeted initial letters should be more likely than standard letters to invoke feelings of the survey being relevant and salient (e.g. [20]).

Motivation is also related to the accuracy and quality of the answers. It is expected that the interviewee will provide more accurate answers when he/she is very motivated [21]. An important indicator of the quality of the data is the amount of item missing data [22]. This is a critical aspect, as analysts typically only use complete cases to perform analyses. However, this results in a loss of information and potentially in biased estimates. [23] investigates the determinants of item non-response, which are underpinned in the disturbances in the question-answer process: understanding the question, retrieval from memory of relevant information, "computation" of a judgement, communication of the answer. In each phase something may go wrong and lead to item non-response. Motivation can affect the question-answer process and thus item non-response [24]. [25] argues that a highly salient question will in many cases improve recall and reduce "do not know" answers.

In the spirit of targeted design, we consider targeted versions of the letter, which aim at emphasising different aspects of the survey content to different people in order to increase the leverage and saliency of the survey. To the best of our knowledge, this kind of manipulation was first considered by [1], where the author investigates the effects of targeted letters on response rates. Building on results in [1], and considering the quality of a survey from a larger perspective, we empirically investigate possible effects of targeted letters with respect to several quality aspects: sample composition, representativity, and item non-response. The analysis thus provides information on how it is possible to improve the quality of the data. A second aspect of interest, which will be the object of another paper, is to study the impact of this method on final estimates. Overall this empirical investigation is important for economic analyses, as it represents an initial step towards the improvement of the panel data quality, that will have an impact for researchers and policy makers using the panel.

The analysis is based on a randomised experiment carried out in the longitudinal study Understanding Society Innovation Panel in 2013 (see Section 3 for detailed description). As pointed out in [26], the longitudinal context entails several advantages for the implementation of targeted designs. Indeed, in longitudinal studies, targeted designs could take advantage of the wealth of information on panel members, collected during recruitment and over time, which could be exploited to improve the data collection process. Some constraints in implementing targeted designs are reduced in the longitudinal framework, e.g. in identifying special groups to be targeted. Nevertheless, such strategies are not used in longitudinal surveys on a standard basis and only a few experiments have been performed [26–28]. Further, the longitudinal context allows a deeper analysis of the effects of the intervention, taking into account several aspects of the sample characteristics.

In Section 2, we formulate our research questions. Section 3 describes the data and the experiment. In Section 4, we describe the results. Section 5 concludes.

2. Research question

[1] investigates whether, and in what circumstances, targeted letters perform better (in terms of response rates) than standard letters. The author does not find an overall increase in response rates. He finds pos-

itive impact for two operationally important low response propensity groups though: previous wave non-respondents and persons who had joined the panel more recently. Further, he finds that the impact differs between modes of data collection. His empirical analysis refers to the same experiment we are considering. Building on the abovementioned results by Lynn, we formulate the following research questions.

First, if targeted letters improve response rates in low response propensity groups, we expect that the composition of the sample changes. Indeed, low response propensity groups are expected to be different from high response propensity groups, with respect to a number of variables. Reluctant respondents and respondents difficult to reach are likely to be quite different [20]. We thus hypothesise that targeted letters change sample composition with respect to standard letters.

Second, interest in the survey topic should increase commitment to respond accurately. Our hypothesis is thus that targeted letters decrease item non-response in the survey questionnaire compared to standard letters.

Finally, given that significant effects on response rates are found only in specific circumstances [1], the third research question is whether the use of targeted letters induces any difference in effects investigated in the first two research questions compared to standard letters under these circumstances.

3. Data and experiment

Understanding Society is a longitudinal panel survey of individuals in the UK [29]. Annual interviews run since 2009 on around 100000 sample members. Surveys are about housing, economic activity, health, income, political attitudes, and several other topics. For methodological experiments and testing that could lead to improvements in the main panel, Understanding Society includes the Innovation Panel (IP). See [30] for details.

In this paper, we focus on Wave 6 of the Innovation Panel (IP6). The target population for the IP is all individuals aged 16 or over and living in England, Scotland, or Wales. The IP consists of a stratified clustered probability sample of households in Great Britain. All members of selected households are included in the panel. An initial sample of 2760 addresses was included from Wave 1 in 2008. 960 were added at Wave 4 in 2011 by means of a refreshment sample. As for data collection modes, in Wave 5 all sample members were randomly allocated to one of two mode treatments [31,32]. One third were allocated to a single-mode CAPI design and two-thirds were allocated to a sequential mixed-mode design in which sample members were first invited to complete the survey online, with non-respondents followed up by CAPI. In Wave 6 the same mode treatment was applied, with an additional telephone mop-up phase.

At Wave 6, a randomized experiment to test whether letters targeted at particular sample subgroups and referring to issues of likely special importance to them could positively affect participation was undertaken [1]. All sample persons aged 16 or over and eligible to be in the field for wave 6 were randomly allocated, with equal probabilities, to two groups (treatment and control group). Random allocation to treatment of standard versus tailored letter was of households within PSUs. The control group received standard letters. The treatment group received targeted initial letters. In the treatment group, five population subgroups, depending on socio-demographic characteristics, were chosen for targeting. These correspond to groups with relatively low propensity to respond. Sample members in the treatment group that was not possible to allocate to any target group due, for example, to missing data, were sent standard letters. This entails six different versions of the letter sent to six different subgroups in the treatment group. Members of the same household may receive different versions of the letter. Targeted groups are defined as follows. Panel members are assigned to group according to a priority order.

– *Employment busy*: employed for at least 39 hours per week, or employed for 30 to 38 hours with a commute of at least 60 minutes
– *With children*: Responsible for at least one child under 15 living in the same household at the time of most recent interview
– *Young*: aged 16 to 29 at the time of wave 5
– *London*: Resident in London or South East England at the time of most recent interview
– *Pensionable*: Of pensionable age at the time of wave 5 (60 or over for women; 65 or over for men)
– *Remainder*: None of the above.

The targeted letter enlisted respondents by saying that Understanding Society "helps researchers and policy makers understand the changes in the needs of the country across diverse subjects like..." and specific group interests were cited. The full letter is available on Understanding Society website.

The analysis presented in this paper is based on 2733 sample persons aged 16 or over that were issued to the

Table 1
Wording in the targeted initial letters and allocations

Group	Text	Interviews
Targeted group	Targeted text	
Employment busy	your work-life balance, your position in your employment and your retirement	223
With children	the provision of child care, schooling and education	173
Young	the impact of the economic climate on employment prospects and the influence of mobile technology on life	175
London	the cost of living and the provision of schools, housing and public transport	185
Pensionable	the provision of social care and the cost of energy and fuel	219
Standard group	Standard text	412

field at wave 6. Control group consists of 1346 units, whereas targeted group consists of 1387 units. Table 1 lists the targeted groups and shows the allocation of the sample and the type of letters for each treatment group. Covariates were measured in Wave 5 or in the last available interview before Wave 5.

In order to better understand the results, it is worth underlying that the IP, being a panel for experiments, has several components where data is collected under different circumstances. With respect to time in the panel, some panel members have been in the panel for six waves (Original sample), others have entered the panel in Wave 4 and Wave 6 is the third wave for them (Refreshment sample). This distinction is important as it is known that attrition in panel surveys is highest in the second wave and then declines over successive waves [33,34].

With respect to mode, data were collected under two different circumstances: mixed-mode (web, followed by CAPI) and CAPI only. Notice that the letter has a different meaning depending on mode. For those assigned to the CAPI group, the initial letter was a prenotification letter sent in advance to the visit by a CAPI interviewer. For the other two-thirds of the sample assigned to the mixed-mode group, the letter was an invitation letter to a web survey. Effects of targeted letters may also differ between the two contexts.

Notably, [1] finds positive impact for two "operationally important" low response propensity groups though: previous wave non-respondents and relatively recent panel entrants. Further, he finds that the impact differs between modes of data collection. In particular, the positive effect on response rates for previous wave non-respondents is restricted to the mixed-mode context, while the positive impact on recent panel members appears to be restricted to single-mode CAPI context.

4. Results

Fieldwork outcome is shown in Table 2, where classification of outcomes is made according to [35]. No-

Table 2
Sample distribution by outcome at wave 6 and treatment group

	Standard	Targeted
Full interview	72.0	73.8
Non-contact	7.5	4.4
Refusal	14.9	14.9
Other non-response	5.6	6.9
N	1346	1387

tice that in the IP, there is a brief proxy interview about adults impossible to be interviewed. We considered these as nonresponse. Since targeted letters were addressed to sample members, it makes little sense to consider their effect on the propensity of other persons to provide proxy response for them. Looking at the reason for proxy response, we classified proxy responses as either noncontact, refusal or other. If the reason was 'Unable to contact', we classified non-response as non-contact, if it was 'Individual refused but allows proxy' we classified non-response as refusal, otherwise we classified non-response as other non-response. This method lead to a rather high level for the category 'other non-response'.

In Section 4.1, we first compare the composition of the panel (respondents) in the two treatment groups, to investigate whether the treatment (targeted letter versus standard letter) produced any difference with respect to sample composition. Next, we look in detail groups where [1] finds significant effects of the treatment on response rates. In Section 4.2, we consider item non-response, both overall and under specific circumstances.

We report unweighted results. This is justified by the fact that we are interested in the functioning of the panel and the impact of the targeted letters on the behaviour of panel members. A weighted analysis inferring to the effects of targeted letters on population distributions and estimates would also be very interesting, especially for substantive users.

4.1. Sample composition

In this section, we examine whether the two different letters have different implications for sample composition. In other words, is sample composition affected by targeted intervention? This issue is very important as it could have an impact on final estimates.

Table 3 shows the marginal distributions in the two treatment groups together with p-values from tests for independence between treatment and the variables gender, age, ethnic group, having a job, household type, number of children in the household, marital status, urbanicity, and education. Overall at the 5 percent level of significance, only age shows a statistically significant difference between the two groups. This is probably related to the fact that the definition of targeted subgroups takes age into account.

Table 4 reports the effect of the targeted letters on variables related to health, the financial situation of panel members and their political interests. This is possible thanks to the wealth of information available on panel members and in general would not be possible in the cross-sectional context. Looking at results, no significant difference is detected between the two samples overall.

Let us now look at specific subgroups where [1] finds differences in response rates: previous wave non-respondents with data collected in a mixed-mode design and recent panel entrants with CAPI interviews. Consider the first group – previous wave non-respondents in mixed-mode design. Tables 3 and 4 show that, for data collected under these circumstances, differences in sample composition may be observed between those who received targeted letters and those who received standard letters. At the 5% level of significance, differences may be observed with respect to age, household type, and a political related variable, namely 'Closer to one political party than others'.

As for the second group where targeted letters are found to be particularly effective – refreshment sample with CAPI interviews – fewer differences in sample composition are observed. At the 5% level of significance, only education shows a statistically significant difference.

In longitudinal surveys it is difficult to ascertain whether the composition of the sample improves after a treatment as there are no reference values to look at. To give an overall idea of the representativeness of the two groups, we compute R-indicators. This set of indicators is based on the concept of representative response. The overall R-indicator is an overall measure of representativity. It varies between 0 and 1, with the value 1 indicating the most representative response and the value 0 indicating the least representative response [36]. Partial indicators provide information on the contribution of single specified variables (and their categories) to a lack of representative response [37]. R-indicators can also be used to derive bias intervals for commonly used estimators [9].

We study the representativity of each group with respect to the sample of individuals issued to the field at Wave 6. We therefore adopt a marginal approach. Refer to [38] for details about studying the representativity of longitudinal panels. To compute the R-indicator, we fitted separate logistic regressions at Wave 6 for each treatment group. The dependent variable is the response indicator. Covariates are socio-demographic variables - gender, age (5 categories), household type, number of children, employment status (2 categories) – as well as mode and time in panel. Predicted response propensities from the model are used to compute the indicators. Adjustments for overall and for variable level indicators are applied (as suggested in [37] for sample sizes less than 15000). The overall R-indicator for the targeted group is 0.82, with confidence interval (0.77, 0.87), while for the control group it is 0.78 (0.71, 0.85). As the two confidence intervals overlap, the difference of the overall representativity in the two groups is not statistically significant. Note though that confidence intervals are quite large and this is due to the relatively small sample sizes in the two groups. Partial indicators show lack of representativity with respect to age and household type. Confidence intervals in the two groups are overlapping.

With respect to the first research question, it is only partially confirmed. Overall, only small changes could be detected in sample composition. As for the third research question related to sample composition, it is confirmed as differences in sample composition could be detected in the two subgroups where improvements in the response rate were observed.

4.2. Item non-response

In this section, we compare the two treatment groups in terms of item non-response. We consider three measures. The first one is an overall item non-response rate. We report the mean item non-response rate, based on over 1000 items in the individual questionnaire, including six items about unearned income sources. The reported rate is the proportion of items for which the respondent was eligible, to which he/she answered

Table 3
Individual response rates, by treatment and characteristics – F2F = face-to-face; MM = mixed modes

Variable	Overall		MM IP5 non-respondents		F2F refreshment	
	Standard	Targeted	Standard	Targeted	Standard	Targeted
Gender (Overall: $P = 0.19$; MM IP5 non-respondents: $P = 0.22$; F2F refreshment: $P = 0.38$)						
Male	44.1	46.1	44.2	53.4	42.5	46.0
Female	55.9	53.9	55.8	46.6	57.5	54.0
Age (Overall: $P = \mathbf{0.08}$; MM IP5 non-respondents: $P = \mathbf{0.02}$; F2F refreshment: $P = 0.48$)						
16–20	5.1	7.4	5.8	16.1	4.7	4.0
21–30	10.9	10.6	20.9	24.6	10.4	13.5
31–40	14.4	13.3	20.9	11.0	16.0	18.3
41–50	16.5	20.9	15.1	22.9	17.0	16.7
51–60	18.4	18.5	17.4	16.1	14.2	22.2
61–70	18.2	16.2	9.3	1.7	23.6	19.0
71+	16.5	13.1	10.5	7.6	15.1	6.3
Ethnic group (Overall: $P = 0.84$; MM IP5 non-respondents: $P = 0.16$; F2F refreshment: $P = 0.95$)						
White British	92.7	92.3	88.7	77.6	92.1	91.8
Other	7.3	7.7	11.3	22.4	7.9	8.2
Has a job (Overall: $P = 0.78$; MM IP5 non-respondents: $P = 0.60$; F2F refreshment: $P = 0.90$)						
No	46.4	45.6	48.0	44.2	46.1	47.2
Yes	53.6	54.4	52.0	55.8	53.9	52.8
Household type (Overall: $P = 0.17$; MM IP5 non-respondents: $P = \mathbf{0.00}$; F2F refreshment: $P = 0.70$)						
Single	15.6	13.6	5.8	4.2	11.2	7.1
Single, kids	3.7	3.5	3.5	1.7	4.7	4.0
Couple	33.5	27.5	29.1	3.4	30.8	29.4
Couple, kids	19.8	21.8	23.3	28.0	29.9	25.4
2+ unrelated	17.2	20.4	25.6	26.3	14.0	24.6
2+ unrelated, kids	10.1	13.1	12.8	36.4	9.4	9.5
Number of own children in household (Overall: $P = 0.91$; MM IP5 non-respondents: $P = 0.68$; F2F refreshment: $P = 0.17$)						
0	75.8	74.9	70.9	64.4	68.2	71.4
1	10.2	10.7	8.1	11.0	10.3	17.5
> 1	13.9	14.4	20.9	24.6	21.5	11.1
Marital status (legal) (Overall: $P = 0.66$; MM IP5 non-respondents: $P = \mathbf{0.07}$; F2F refreshment: $P = 0.99$)						
Single and never married	18.7	20.4	19.5	34.6	16.3	15.9
Relationship	65.8	64.8	64.6	55.4	73.1	73.8
Separated/Divorced/Widowed	15.5	14.8	18.9	10.0	10.6	10.3
Urban/Rural (Overall: $P = 0.45$; MM IP5 non-respondents: $P = 0.55$; F2F refreshment: $P = 0.12$)						
Rural	25.3	23.3	29.1	23.7	30.8	17.5
Urban	74.7	76.7	70.9	76.3	69.2	82.5
Education (Overall: $P = 0.61$; MM IP5 non-respondents: $P = 0.89$; F2F refreshment: $P = \mathbf{0.02}$)						
Degree	20.7	20.0	16.2	17.0	24.5	15.5
Other higher degree	12.7	10.4	8.1	9.6	15.7	8.9
A-level etc.	19.1	21.7	25.7	22.3	18.6	31.7
GCSE etc.	21.8	23.3	27.0	25.5	19.6	27.6
Other qualification	13.2	12.4	9.5	14.9	10.8	3.2
No qualification	12.5	12.2	13.5	10.6	10.8	13.0

"don't know" or "refused". The second measure refers to the proportion of employed respondents who did not provide an answer for their last gross pay. This item is important for the derivation of income measures and it is prone to high levels of item non-response. The third measure is the proportion of eligible respondents who did not provide information on savings and investments. This item was selected as it is the one with highest levels of item non-response in the questionnaire.

Table 5 reports results for the overall sample and for specific subgroups. We see that overall no differences are observed between the two groups with respect to item non-response. Looking at specific subgroups, we see that for previous wave non-respondents with data collected by mixed-mode, the item non-response rate for last gross pay was significantly higher with targeted letter than standard letter ($P = 0.04$). Thus, for this group, higher response rates were observed with targeted letters but the quality of the answers, as measured by item non-response, is lower.

Looking at the refreshment sample with CAPI interviews, at the 10% level of significance a statistically significant difference is observed for item non-response rate for last gross pay. Targeted letters show lower level of item non-response rate.

The second hypothesis is thus rejected: no differ-

Table 4
Individual response rates, by treatment and characteristics – F2F = face-to-face; MM = mixed modes

Variable	Overall		MM IP5 non-respondents		F2F refreshment	
	Standard	Targeted	Standard	Targeted	Standard	Targeted
General Health (Overall: $P = 0.88$; MM IP5 non-respondents: $P = 0.63$; F2F refreshment: $P = 0.26$)						
Excellent/Very good	45.7	44.8	45.3	51.6	51.0	41.5
Good	30.4	31.6	28.0	22.1	28.4	36.6
Fair/Poor	23.9	23.7	26.7	26.3	20.6	21.9
Savings and investments (Overall: $P = 0.97$; MM IP5 non-respondents: $P = 0.48$; F2F refreshment: $P = 0.21$)						
No	66.8	66.7	74.1	79.3	63.4	71.9
Yes	33.2	33.3	25.9	20.7	36.6	28.1
How managing financially now (Overall: $P = 0.70$; MM IP5 non-respondents: $P = 0.23$; F2F refreshment: $P = 0.95$)						
Living comfortably	25.9	26.8	13.3	24.7	29.7	31.1
Doing alright	36.7	34.6	26.7	25.9	37.6	35.3
Just/difficult	37.4	38.6	60.0	49.4	32.7	33.6
How manage financially in future (Overall: $P = 0.81$; MM IP5 non-respondents: $P = \mathbf{0.006}$; F2F refreshment: $P = 0.12$)						
Better off	18.5	18.9	26.7	30.9	19.2	16.4
Worse off than you are now	22.1	23.2	31.7	16.0	17.2	30.3
About the same	59.4	57.9	41.6	53.1	63.6	53.3
Supports a particular party (Overall: $P = 0.55$; MM IP5 non-respondents: $P = 0.39$; F2F refreshment: $P = 1.00$)						
Yes	32.9	31.5	35.0	28.4	28.7	28.7
No	67.1	68.5	65.0	71.6	71.3	71.3
Closer to one political party than others (Overall: $P = 0.46$; MM IP5 non-respondents: $P = \mathbf{0.04}$; F2F refreshment: $P = 0.41$)						
Yes	36.2	34.1	41.7	23.2	29.1	35.5
No	63.8	65.9	58.3	76.8	70.9	64.5
Level of interest in politics (Overall: $P = 0.88$; MM IP5 non-respondents: $P = 0.15$; F2F refreshment: $P = 0.68$)						
Very/Fairly	46.2	46.7	41.7	54.3	49.5	46.7
Not very/Not at all	53.8	53.3	58.3	45.7	50.5	53.3

Table 5
Item non-response (INR) rates for the overall sample and specific subgroups, by treatment – F2F = face-to-face; MM = mixed modes

	Overall			MM IP5 non-respondents			F2F refreshment		
	Std	Targeted	P	Std	Targeted	P	Std	Targeted	P
Overall Mean INR (%)	1.14	1.32	0.28	1.46	2.58	0.11	52.6	45.3	0.48
N	984	1040		86	118		109	131	
Variable									
Gross pay: INR(%)	14.4	14.8	0.90	17.5	38.3	**0.04**	16.7	3.2	**0.07**
N	472	507		40	60		48	62	
Savings and investments: INR (%)	37.3	36.4	0.78	43.0	52.5	0.19	17.4	11.4	0.26
N	982	1040		86	118		109	131	

ences are found with respect to item non-response between the two groups overall. As for the third hypothesis, it is partially accepted: CAPI interviews in the refreshment are more accurate when targeted letters are received, the opposite happens in mixed-mode for previous wave non-respondents.

5. Conclusions

Motivating in survey participation is an important issue for survey quality both in general population and in business surveys. The matter is relevant for official statistics as well as for academic and market research surveys. Understanding how motivation affects participation could help official statistics to take advantage of the integration of the survey process with administrative data and paradata to redesign the surveys and to monitor the survey process.

In this study, we have discussed some evidence on the effects of targeted initial letters in a general population panel, with respect to sample composition and item non-response.

The findings show that targeted initial letters may have an effect on sample composition. The effects are uneven across sample subgroups and survey design circumstances. Effects are found also with respect to item non-response.

We particularly investigated cases where improvements in response rates were observed [1], namely, previous wave non-respondents under mixed-mode design and recent panel members with CAPI interviews.

For the first group – previous wave non-respondents under mixed-mode design – differences are found for sample composition with respect to a number of variables. As it regards the quality of the answers, we found higher levels of item non-response for last gross pay. As for the second subgroup identified by [1] – recent panel members with CAPI interviews – we find fewer differences in sample composition between the two treatment groups. Item non-response rate for last gross pay is lower in the targeted group in this case.

This finding suggests that targeted letters may give a useful contribution in the longitudinal context for the maintenance of the panel. However, results are not conclusive and appear to be dependent on specific contents. Further experiments and studies should be undertaken to better clarify the effect of targeted letters.

Acknowledgments

References

[1] P. Lynn, Targeted appeals for participation in letters to panel survey members. *Public Opinion Quarterly* **80**(3) 2016, 771–782.

[2] R.M. Groves and S.G. Heeringa, Responsive design for household surveys: Tools for actively controlling survey errors and costs, *Journal of the Royal Statistical Society: Series A* **169** (2006), 439–457.

[3] J. Wagner, Adaptive survey design to reduce non-response bias. PhD Thesis, University of Michigan 2008.

[4] A. Luiten and B. Schouten, Tailored fieldwork design to increase representative household survey response: An experiment in the Survey of Consumer Satisfaction, *Journal of the Royal Statistical Society: Series A* **176** (2013), 169–189.

[5] D. Massey and R. Tourangeau, New challenges to social measurement, *The Annals of the American Academy of Political and Social Science* **645** (2013), 6–22.

[6] R. Tourangeau, M. Brick, S. Lohr and J. Li, Adaptive and responsive survey designs: a review and assessment, *Journal of the Royal Statistical Society, Series A* (2016), 1–21.

[7] P. Lundquist and C.E. Särndal, Aspects of responsive design with applications to the Swedish Living Conditions Survey, *Journal of Official Statistics* **29** (2013), 557–582.

[8] N. Shlomo, B. Schouten and V. de Heij, Designing Adaptive Survey Designs with R-Indicators. Paper presented at the New Techniques and Technologies for Statistics Conference 2013, Brussels.

[9] B. Schouten, F. Cobben, P. Lundquist and J. Wagner, Does more balanced survey response imply less non-response bias? *Journal of the Royal Statistical Society Series A* **179**(3) (2016), 727–748.

[10] D.A. Dillman, J.D. Smyth and L.M. Christian, Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method. 4th ed. Hoboken, NJ: John Wiley, 2014.

[11] R.M. Groves and M.P. Couper, Nonresponse in Household Interview Surveys. New York: Wiley, 1998.

[12] E. Singer, Exploring the Meaning of Consent: Participation in Research and Beliefs about Risks and Benefits, *Journal of Official Statistics* **19** (2003), 273–285.

[13] R.M. Groves, R.B. Cialdini and M.P. Couper, Understanding the decision to participate in a survey, *Public Opinion Quarterly* **56**(4) (1992), 475–495.

[14] M.P. Couper, E. Singer, F.G. Conrad and R.M. Groves, Risk of Disclosure, Perceptions of Risk, and Concerns about Privacy and Confidentiality as Factors in Survey Participation, *Journal of Official Statistics* **24**(2) (2008), 255–275.

[15] M. Kropf and J. Blair, Testing Theories of Survey Cooperation: Incentives, Self-Interest and Norms of Cooperation, *Evaluation Review* **29**(6) (2005), 559–575.

[16] C.L. Martin, The impact of topic interest on mail survey response behaviour, *Journal of Market Research Society* **36** (1994), 327–338.

[17] H. Roose, H. Waege and F. Agneessens, Respondent related correlates of response behavior in audience research, *Quality & Quantity* **37** (2003), 411–434.

[18] H. Roose, J. Lievens and H. Waege, The joint effect of topic interest and follow-up procedures on the response in a mail questionnaire, *Sociological Methods & Research* **35** (2007), 410–428.

[19] R.M. Groves, E. Singer and A. Corning, Leverage-saliency theory of survey participation: description and an illustration, *Public Opinion Quarterly* **64** (2000), 299–308.

[20] R.M. Groves and M. Couper, Nonresponse in Household Interview Surveys. New York: Wiley, 1998.

[21] C.F. Cannell, P.V. Miller and L. Oksenberg, Research on Interviewing Techniques, *Sociological Methodology* **12** (1981), 389–437.

[22] R.M. Groves, Survey Errors and Survey Costs. New York: Wiley, 1989.

[23] E.D. De Leeuw, J. Hox and M. Huisman, Prevention and treatment of item nonresponse, *Journal of Official Statistics* **19** (2003), 153–176.

[24] P. Beatty and D. Herrmann, To answer or not to answer: Decision processes related to survey item nonresponse. In: R. Groves, D. Dillman, J. Eltinge, R. Little, eds, Survey nonresponse. New York: John Wiley & Sons, 2002, pp. 71–69.

[25] N. Mathiowetz, Presentation in Response to T. Murato and P.A. Gwartney, Question Saliency, Question Difficulty, and Item Nonresponse in Survey Research. International Conference on Survey Nonresponse, Portland, October 1999.

[26] P. Lynn, Targeted response inducement strategies on longitudinal surveys. In: U. Engel, B. Jann, P. Lynn, A. Scherpenzeel and P. Sturgis, eds, Improving Survey Methods: Lessons from Recent Research. New York: Routledge/Psychology Press, 2015, pp. 322–338.

[27] A. Bianchi and S. Biffignandi, Responsive design for economic data in mixed-mode panels. In: Mecatti F, Conti PL, Ranalli MG, eds, Contribution to Sampling Statistics. Springer, 2014, pp. 85–102.

[28] L. Calderwood, A. Cleary, G. Flore and R.D. Wiggins, 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 2012.

[29] N. Buck and S. McFall, Understanding Society: Design Overview, *Longitudinal and Life Couse Studies* **31** (2012), 5–17.

[30] Uhrig SCN. Using experiments to guide decision making in Understanding Society: Introducing the Innovation Panel, chapter 13. In: S.L. McFall and C. Garrington, eds, Understanding Society: Early Findings from the First Wave of the UK's Household Longitudinal Study. Colchester: University of Essex, 2011. Available from: http://research. understandingsociety.org.uk/findings/early-findings.

[31] A. Jäckle, P. Lynn and J. Burton, Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response, *Survey Research Methods* **9** (2015), 57–70.

[32] A. Bianchi, S. Biffignandi and P. Lynn, Web-CAPI sequential mixed mode design in a longitudinal survey: effects on participation rates, sample composition and costs. To appear in Journal of Official Statistics 2017.

[33] P. Lugtig, Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers, *Sociological Methods and Research* **43** (2014), 699–723.

[34] R.F. Schoeni, F. Stafford, K.A. McGonagle and P. Andreski, Response Rates in National Panel Surveys, *Annals of the American Academy of Political and Social Science* **645** (2013), 60–87.

[35] American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 8th ed., Lenexa: American Association for Public Opinion Research, 2015.

[36] B. Schouten, F. Cobben and J. Bethlehem, Indicators for the representativeness of survey response. Survey, *Methodology* **35** (2009), 101–113.

[37] B. Schouten, N. Shlomo and C. Skinner, Indicators for Monitoring and Improving Representativeness of Response, *Journal of Official Statistics* **27** (2011), 231–253.

[38] A. Bianchi and S. Biffignandi, Representativeness in panel surveys. To appear in Mathematical Population Studies (2017).