

# Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response

Annette Jäckle

Institute for Social and Economic Research  
University of Essex

Peter Lynn

Institute for Social and Economic Research  
University of Essex

Jonathan Burton

Institute for Social and Economic Research  
University of Essex

There are considerable cost and timeliness advantages associated with web interviewing, compared to interviewer administration. However, web surveys do not perform well in terms of coverage and participation. To harness the strengths of both modes, existing probability-based interviewer-administered surveys are therefore being pushed to consider a mixed mode approach, including web. We assess the effect of introducing web interviewing as part of a mixed-mode design in the context of an existing longitudinal survey in which sample members have previously been interviewed face-to-face. Using experimental data from a household panel survey in the UK, we find that the mixed mode design resulted in a lower proportion of households fully responding. However, more than one in five households fully responded online. Overall, individual response rates were also lower with the mixed mode design, and we were unable to identify any subgroups where the reverse was true. Also, item nonresponse rates were higher with the mixed mode design.

*Keywords:* longitudinal survey; mixed mode survey; non-response; respondent incentives; response rate; web survey

## 1 Introduction: Mixed-Mode Survey Data Collection

Considerable attention is being given to survey data collection designs which mix interviewer administration with web. The incorporation of web into a mixed mode design has potential both to reduce survey costs and improve quality (Couper, 2011; Groves & Lyberg, 2010; Kreuter, 2013). Several UK government surveys, which currently involve face-to-face interviewing, are actively considering mixed mode approaches (Betts & Lound, 2010). A Europe-wide project is considering ways of incorporating web data collection into the European Labour Force Surveys (Luiten & Blanke, 2013), which tend to currently rely on either computer-assisted telephone interviewing (CATI) or a mix of CATI and face-to-face computer-assisted personal interviewing (CAPI).

However, experience of implementing mixed mode surveys which include both interviewer administration and web is severely limited<sup>1</sup>, especially in the context of academic or

government longitudinal surveys. Aside from specialist and commercial applications, few major surveys have used mixed mode designs including web, and only in limited ways (Dex & Gumy, 2011). Though some surveys have experimented with web data collection, they have not yet implemented it on the production survey. For example, Griffin, Fischer, and Morgan (2001) report an experiment on the American Community Survey (ACS) in which web was offered as an additional mode in an otherwise mail-CATI-CAPI sequential mixed-mode design. The inclusion of web reduced the overall response rate and it was unclear whether cost savings would accrue, so it was decided not to include web in the ACS design. Similarly, Lagerstrøm (2011) reports experiments on the Norwegian Rent Market Survey which highlighted concerns about differential measurement and did not indicate a likely cost saving. The most encouraging experiments to date are those carried out by Statistics Netherlands (Cuppen, Van Der Laan, & Van Nunspeet, 2011), in which response rates with a web-CATI-CAPI design were no lower

---

Annette Jäckle, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK (aejack@essex.ac.uk)

<sup>1</sup> By “mixed mode” we mean surveys that collect the *same* data from different respondents in different modes. This is distinct from the more common “multiple mode” design, in which *different* data are collected in different modes, e.g. a self-completion component within a face-to-face interview.

than with a CATI-CAPI design, and for which it was concluded that cost savings were associated with the design including web. These studies all differ in context from ours in a couple of important respects. First, they involved contacting fresh samples of individuals or households, whereas we are concerned with re-contacting panel survey respondents. Second, they all include CATI as part of the mode mix, whereas we are concerned with a mix of web and CAPI.

The longitudinal survey context introduces some specific issues. First, response rates are arguably more important than in a cross-sectional survey. Over and above concerns about non-response bias, high response rates, at least from wave 2 onwards, are essential in order to maintain the sample available for longitudinal analysis. Non-responding sample members cannot be replaced by adding new samples as it would not be possible to collect data equivalent to those already collected at previous waves. Second, in the case of an existing survey, sample members will have prior experience of being interviewed in another mode and prior knowledge of the survey content. Conceivably, this familiarity might improve the chances of response in the absence of interviewer explanation and persuasion. Also, the respondents will have already been recruited in another mode, another feature which could make it easier to persuade them to take part by a self-completion mode. As these features may affect response propensity, it is important to test the effects of mixed mode designs on response rates in a longitudinal context, as we do in this study.

A large literature attests to the effectiveness of respondent incentives in improving response rates, for both interviewer-administered surveys (Cantor, O'Hare, & O'Connor, 2008; Singer, Gebler, Raghunathan, Van Hoewyk, & McGonagle, 1999) and mail surveys (Church, 1993; Singer & Ye, 2013), especially when the incentive is monetary in form and offered unconditionally in advance (Church, 1993; Singer et al., 1999; Singer & Ye, 2013). The size of the effect generally appears to be larger for mail surveys (Singer & Ye, 2013), offering hope that web surveys, another form of self-completion survey, could also benefit considerably from incentives. The literature on the effects of incentives on web surveys is smaller, but growing, and also suggests a generally positive effect (Göriz, 2006, 2010, 2015), though lotteries rather than monetary incentives are often preferred, for reasons of both cost and practicality. Incentives seem to be at least as effective in longitudinal surveys as they are in cross-sectional surveys (Laurie & Lynn, 2009; Schoeni, Stafford, McGonagle, & Andreski, 2013), but very few studies of incentives have been carried out in a mixed mode longitudinal context (Jäckle & Lynn, 2008) and none, to our knowledge, in the context of introducing a mixed mode design for the first time.

There are few direct comparisons of item nonresponse rates in web and CAPI. However, we expect that item non-

response rates will tend to be higher in web unless the combination of the nature of the question and the design of the web instrument is particularly favourable, in which case rates may approach those achieved in CAPI. This expectation is a corollary of the conclusion of Tourangeau, Conrad, and Couper (2013, p. 54) that "... when web surveys are designed to resemble paper surveys ... the missing data rates will resemble those of mail surveys, but that when the web surveys are designed to take advantage of the interactive capabilities of the web, missing data rates will be lower;" in combination with earlier research showing that item nonresponse rates tend to be higher in (paper) self-completion surveys than either CAPI (Bowling, 2005; De Leeuw, 2005; Nicolaas, Thomson, & Lynn, 2000) or CATI (De Leeuw, 2005; Fricker, Galesic, Tourangeau, & Yan, 2005; Harris, Weinberger, & Tierney, 1997). However, the longitudinal context again introduces distinct considerations. For example, respondents who might otherwise fear certain repercussions as a result of giving a particular response should have observed from previous waves the absence of such repercussions. Furthermore, if panel sample members have previously completed very similar interviews with face-to-face interviewers they may now have better comprehension of the questions than they would have had in the absence of any previous interviewer involvement. Thus, the specific context of a longitudinal survey in which previous waves were interviewer-administered is unique. We are unaware of previous research into item non-response in this context.

Possible effects of measurement error are an important concern when considering the introduction of a mixed mode design in any context (Bowling, 2005), including when the mix consists of web and face-to-face (Duffy, Smith, Terhanian, & Bremer, 2005). In the longitudinal context, the possibility of individuals responding in different modes at different waves introduces concerns about measurement error affecting within-unit measures of change (Dillman, 2005, 2009), as well as affecting between-unit differences, which is the usual concern with cross-sectional surveys. However, measurement error is not the focus of this article. We focus here on unit and item nonresponse.

Similarly, data collection costs are not the focus of our study, though they do provide a motivating context for our experiment. A meaningful comparative analysis of costs would require careful consideration of all the factors affecting between-mode cost ratios, such as sample size, sample design, interview length, interviewer fee structure, survey design and various constraints. This substantial work should be the subject of a separate paper.

We assess the effects of introducing a sequential mixed mode design, involving both web interviewing and face-to-face interviewing, in the context of an ongoing longitudinal survey in which sample members have previously been interviewed face-to-face. An important feature of the longitu-

dinal survey in question is that at each wave interviews are required not only with each sample person but also with all other adult members of their household. This is a feature common to most household panel surveys. In this context, a significant cost-saving is associated with the introduction of the mixed-mode design only when all household members respond by web. If one or more household member does not respond by web, it is necessary to send a face-to-face interviewer to visit the household to seek the remaining interview(s). In so far as the main motivation for introducing web interviewing is to save costs, a central interest is therefore the extent to which whole-household participation by web is possible. This is a demanding target and may make cost savings harder to achieve than in a simple survey of individuals. Moreover, the importance of response rates, coupled with a recognition that co-operation may be harder to maintain in the absence of personal interviewer contact, means that we must simultaneously be concerned with overall response rates achieved by the mixed mode design, regardless of the mode of response.

In the next section we outline the context and design of our experimental study. Following sections present results regarding the extent of full household response online (section 3), unit response rates and sample composition (section 4) and item response rates (section 5). Section 6 discusses the implications, both for household panel surveys and for mixed mode surveys more generally, as well as suggesting some avenues for future research.

## 2 Study Design

Our study was carried out at wave 5 of the *Understanding Society* Innovation Panel (Uhrig, 2011), in 2012, and was designed to help establish whether, and in what circumstances, it may be appropriate for the main *Understanding Society* survey (Buck & McFall, 2012) to move to a design that includes web interviewing<sup>2</sup>. Like the main survey, the *Understanding Society* Innovation Panel involves interviews at 12-month intervals with a sample of persons and with all other members of the current household of each sample person. The interviews have broad topic coverage (household dynamics, income and welfare, labour market participation, health and wellbeing, political attitudes, social participation). The sample had two components:

- Original sample, for whom this was the 5th wave (1,090 households; 2,112 adults eligible for interview);
- Refreshment sample, for whom this was the 2nd wave (483 households; 928 adults).

Both samples are stratified, clustered, probability samples of all persons resident in Great Britain (excluding the northernmost part of Scotland) at the time of the first wave of field work for that sample (spring 2008 for the original sample, and summer 2011 for the refreshment sample). Primarily sampling units are postal sectors, geographical areas

that contain around 2,500 households on average. Secondary sampling units are residential addresses selected from the Postcode Address File (Lynn & Lievesley, 1991) and sample elements are persons. The sample of persons is therefore initially clustered within households, but this clustering reduces over time as household composition changes. More details of the Innovation Panel sample design can be found in Lynn (2009). At each of waves 1 to 4, sample members aged 16 or over were interviewed in person; those aged 10–15 were administered a paper self-completion questionnaire; and information about those aged 0–9 was collected from a parent or other responsible adult. At waves 1, 3 and 4, interviews were carried out solely by CAPI. Experimentation with a mixture of CAPI and computer-assisted telephone interviewing (CATI) was carried out at wave 2 of the Innovation Panel in 2009 (Lynn, Uhrig, & Burton, 2010). The main conclusion from that experiment was that a CATI-CAPI sequential mixed modes design, if implemented in a way that would save costs, was likely to result in lower response rates (Lynn, 2013). For that reason, CATI was not included in the wave 5 experiment.

Only sample members who were in participating households at the first wave for that sample were re-approached for interview at each subsequent wave. From wave 2 onwards, non-response at one wave did not preclude an interview attempt at the next wave. Thus, in our study (at wave 5), the original sample included individuals that had responded at wave 4 and a number of individuals and whole households that had not been contacted or had refused at wave 4. The Refreshment sample only included individuals in households that had responded at wave 4. Households in both samples were randomly assigned to one of two treatment groups, so all individuals in the same household received the same treatment:

- Face-to-face (one-third of each sample);
- Mixed mode (two-thirds of each sample).

The distribution of the issued sample of households across samples and mode treatments is summarised in Table 1. The randomisation was implemented within primary sampling units (PSUs), so that each PSU – which typically formed one interviewer assignment – contained a mix of households in both treatment groups.

The face-to-face treatment involved standard *Understanding Society* procedures. Each adult sample member (aged 16 or over) was sent an advance letter with an unconditional incentive, after which interviewers visited to attempt CAPI interviews. The value of the incentive (in both samples) was subject to experimental allocation. Original sample members received either £5 or £10, while refreshment sample members received £10, £20 or £30. As with the modes ex-

<sup>2</sup> For a review of the risks and opportunities for the *Understanding Society* survey associated with a move to a mixed mode design involving web, see Couper (2012).

periment, allocation was at the household level, so all individuals in the same household received the same incentive. The mode treatments and the incentive treatments were fully crossed. In each household one person was asked to complete the household enumeration grid and household questionnaire, an interview that averaged around 12 minutes. All household members aged 16 or over were asked for an individual interview (average around 35 minutes) and to complete a self-completion questionnaire, which was randomly allocated to be either a Computer-Assisted Self Interview (CASI) or a paper questionnaire booklet. Young people aged 10-15 were administered a paper self-completion questionnaire.

The mixed mode treatment was as follows. Sample members aged 16 or over were sent a letter with the unconditional incentive, inviting them to take part by web. The letter (see the appendix) included the URL and a unique user ID, which was to be entered on the welcome screen. A version of the letter was additionally sent by email to all sample members for whom we had an email address (around one-third of the sample). For people who had indicated at previous waves that they do not use the internet regularly for personal use, the letter mentioned that they would also have the opportunity to do the survey with an interviewer. Up to two email reminders were sent at three-day intervals. Sample members who had not completed the web interview after two weeks were sent a reminder by post and interviewers then started visiting them to carry out CAPI interviews. The interviewer visits began in the same week that the reminder letter would have been received in order to constrain the overall field work period to one that would be feasible on the main survey, which is divided into monthly samples. The web survey remained open throughout the fieldwork period.

The first household member to log on to do the web survey was asked to complete the household grid, which collects information on who is currently living in the household. The web grid included an additional question to identify who is responsible for paying bills. The household questionnaire could be completed by either this person or their spouse/partner. For these sample members the household questionnaire was displayed first, then leading on to the individual questionnaire. (The household questionnaire is relatively short, at around 10 minutes, and collects household-level information such as housing tenure, rent/mortgage payments, expenditure, utility bills, household consumer durables and some measures of material deprivation.) Once one partner had completed the household questionnaire, it would not appear for the other partner.

The web questionnaire was based on the CAPI one, with some adaptations, e.g. incorporating interviewer instructions into question wording, removing references to showcards, and making “help” screens more respondent-appropriate. To replicate the CAPI approach, in which “don’t know” and

“refuse” are not suggested explicitly as options to respondents, these options were not included in the web instrument. However, when a respondent clicked “next” without indicating a response to a question, the question re-appeared on screen but with the “Don’t Know” and “Refuse” options appearing for the first time. There was a message on-screen simply asking the respondent to select an option in order to move on. The web survey was not suitable for completion using a small mobile device (e.g. smartphone). If a mobile device was used to access the log-on page, the respondent was automatically directed to a page requesting that they log on from a computer.

The questionnaire for 10–15 year-olds was administered either on paper or by web, depending on the mode used by the parent(s). If the parent(s) had responded by web and we had their email address, an invitation was sent to the parent by email with a request to forward it to their child. If the parent had been interviewed in CAPI, the interviewer handed out the youth paper self-completion questionnaire. Other youth questionnaires were sent by post.

The mixed mode treatment also included two randomised experiments (fully crossed with each other and with the mode and unconditional incentive experiments) to test ways of increasing web response rates:

- Half the households were offered an additional conditional incentive: if all eligible household members completed the web survey within two weeks, they each received an additional £5. This was mentioned in the advance letters to all household members in this treatment group.
- Half the households were sent the advance letter and first email to arrive on a Friday. The other half were sent to arrive on a Monday.

### 3 Complete Household Response by Web

Significant cost savings can only be made when the need to send an interviewer to visit a household is avoided. This requires that the household responds completely by web (grid, household questionnaire and all individual questionnaires for adults).

Table 2 shows, for each sample and each level of incentive, the proportion of households that completed all instruments by web. This proportion depends on the level of unconditional incentive offered. Higher unconditional incentive levels significantly and substantially increased the probability of participating fully by web for the refreshment sample: with a £10 unconditional incentive per person, 22.2% of households completed all instruments by web; with £20 the figure was 37.8% and with £30 it was 43.0% ( $P = 0.00$ ). For previous-wave responding households in the original sample, the association between the probability of participating fully by web and the incentive level was only marginally significant ( $P = 0.08$ : 18.9% with £5 incentive, 24.4% with £10). This may be because of the comparatively low levels of the

Table 1  
Allocation of Households to Experimental Groups in Wave 5

|              | Original Sample     |                           | Refreshment sample | Total |
|--------------|---------------------|---------------------------|--------------------|-------|
|              | Responded at wave 4 | Did not respond at wave 4 |                    |       |
| Face-to-face | 321                 | 43                        | 168                | 532   |
| Mixed modes  | 615                 | 111                       | 315                | 1041  |
| Total        | 936                 | 154                       | 483                | 1573  |

1503 households were issued to the field, based on information held prior to the start of fieldwork. During the course of fieldwork, 70 additional (split) households were identified that were eligible for interviews. In the Refreshment sample, only responding households from their first wave (wave 4 of the panel) were issued at wave 5.

incentive or because sample members had been in the study for a longer period of time and so were more accustomed to a face-to-face approach. There was no difference in rates of complete participation between the original and refreshment samples, after controlling for incentive level and response at the previous wave. Amongst previous wave respondents given a £10 incentive (the only treatment that was in common across the samples), the proportion of households fully responding by web was 24.4% in the original sample and 22.2% in the refreshment sample. Thus, with this level of incentive nearly one in four sample households did not require an interviewer visit.

A further indicator of potential cost savings is provided by the mean number of interviewer visits per household. This was lower with the mixed modes treatment than with face-to-face, both in the original sample (2.9 vs. 3.7,  $P = 0.00$ ) and in the refreshment sample (2.5 vs. 3.7,  $P = 0.00$ ). It does not seem to be the case that only the “easy” households participated by web. Had that been the case, we should have observed larger numbers of interviewer visits amongst the households in the mixed modes treatment group that remained to be contacted by interviewer after others had participated by web. Further support for this interpretation is provided by an analysis of the effort needed at the previous wave. Using the number of calls an interviewer had to make at the previous wave as a measure of the effort required, those households at wave 5 that completed by web had required 3.5 calls at wave 4, compared to 3.6 calls for those wave 5 households that had not completed by web ( $P = 0.59$ ).

Full household response by web was considerably more likely if the household had internet access at home, if some or all household members were regular web users and if all household members had previously supplied an email address (Table 3). The number of persons aged 16 or over in the household was not associated with the probability of full household response by web in the refreshment sample, but in the original sample there was a higher probability for households with two or more adults to respond fully by web than

households with just one adult.

Furthermore, the proportion of households participating fully by web could perhaps be increased further by careful use of other design features. In Table 4 we summarise the results of the two experiments with design features that might increase web take-up. The conditional incentive increased the proportion of households fully responding by web from 15.6% to 22.3% in the original sample, and from 17.4% to 25.2% amongst previous wave responding households. Although not shown in the table, the effect may have been stronger amongst households in which sample members received the £10 unconditional incentive (16.5% vs. 25.8%;  $P = 0.03$ ) than amongst households in which sample members received the £5 unconditional incentive (14.9% vs. 19.6%;  $P = 0.18$ ). There appears to be no effect of the conditional incentive on the refreshment sample. For all samples the proportion of fully responding households was higher if the web invitation was received on a Friday rather than a Monday, but none of these differences reached statistical significance at the 0.05 level. That was also true for the subset of households who received the invitation by email (result not shown).

#### 4 Unit Response Rates

The impact of mixed-mode data collection on response rates is a further crucial consideration. Here we consider household response rates, individual response rates, and subgroup differences in response rates.

##### 4.1 Household Response Rates

The proportion of households participating did not differ significantly between treatments for either the original or refreshment sample (Table 5). However, amongst previous-wave responding households in the original sample the mixed mode design resulted in fewer complete households (household questionnaire and *all* individual interviews

Table 2  
*Proportion of Households Fully Responding by Web, by Sample and Unconditional Incentive*

|                         | Original sample |     |                              |     |                                  |     | Refreshment sample |     |
|-------------------------|-----------------|-----|------------------------------|-----|----------------------------------|-----|--------------------|-----|
|                         | Total sample    |     | Wave 4 respondent households |     | Wave 4 non-respondent households |     | Total sample       |     |
|                         | %               | n   | %                            | n   | %                                | n   | %                  | n   |
| Total sample            | 18.9            | 726 | 21.3                         | 615 | 5.4                              | 111 | 35.2               | 315 |
| Unconditional incentive |                 |     |                              |     |                                  |     |                    |     |
| £5                      | 17.3            | 399 | 18.9                         | 339 | 8.3                              | 60  | -                  | -   |
| £10                     | 21.1            | 323 | 24.4                         | 275 | 2.1                              | 48  | 22.2               | 90  |
| £20                     | -               | -   | -                            | -   | -                                | -   | 37.8               | 111 |
| £30                     | -               | -   | -                            | -   | -                                | -   | 43.0               | 114 |
| P                       | 0.18            |     | 0.08                         |     | 0.15                             |     | 0.00               |     |

P-values from Wald tests of the equality of mean completion rates between incentive groups, adjusted for sample design.

Table 3  
*Proportion of Households Fully Responding by Web, by Household Characteristics*

| Prob(complete HH by web)  | Original sample (total) |         | Original sample (IP4 respondents) |         | Refreshment sample    |         |
|---------------------------|-------------------------|---------|-----------------------------------|---------|-----------------------|---------|
|                           | Predicted probability   | P-value | Predicted probability             | P-value | Predicted probability | P-value |
| HH internet               |                         |         |                                   |         |                       |         |
| No                        | 8.5                     |         | 6.5                               |         | 22.0                  |         |
| Yes                       | 20.1                    | 0.03    | 22.4                              | 0.00    | 36.7                  | 0.09    |
| Web users in HH           |                         |         |                                   |         |                       |         |
| None <sup>a</sup>         | 7.3                     |         | 8.4                               |         | 8.6                   |         |
| Some                      | 17.7                    |         | 20.4                              |         | 36.2                  |         |
| All                       | 27.5                    |         | 28.7                              |         | 45.1                  |         |
| Yes+missing               | 12.6                    | 0.00    | 14.6                              | 0.00    | 33.0                  | 0.00    |
| Number of eligible adults |                         |         |                                   |         |                       |         |
| 1                         | 14.9                    |         | 18.4                              |         | 30.6                  |         |
| 2                         | 20.1                    |         | 21.5                              |         | 39.5                  |         |
| 3+                        | 24.3                    | 0.07    | 24.8                              | 0.47    | 29.6                  | 0.15    |
| Emails for HH members     |                         |         |                                   |         |                       |         |
| None                      | 8.8                     |         | 9.4                               |         | 25.4                  |         |
| Some                      | 14.7                    |         | 16.5                              |         | 30.4                  |         |
| All                       | 27.9                    | 0.00    | 30.5                              | 0.00    | 48.2                  | 0.02    |
| N                         | 707                     |         | 612                               |         | 314                   |         |

Predicted probabilities from logit model. Estimates adjusted for sampling design. Sample size for IP4 non-respondents too small for estimation.

<sup>a</sup> includes no+missing; no+yes+missing.

Table 4

*Effects of Conditional Incentive and Day of Mailing on Proportion of Households Who Completed All Interviewing by Web*

|                          | Original sample |      |     |                |      |     |                    |      |    |                    |      |     |
|--------------------------|-----------------|------|-----|----------------|------|-----|--------------------|------|----|--------------------|------|-----|
|                          | Total           |      |     | IP4 responding |      |     | IP4 non-responding |      |    | Refreshment sample |      |     |
|                          | %               | P    | N   | %              | P    | N   | %                  | P    | N  | %                  | P    | N   |
| No conditional incentive | 15.6            | –    | 359 | 17.4           | –    | 304 | 5.5                | –    | 55 | 34.8               | –    | 158 |
| Conditional incentive    | 22.3            | 0.01 | 363 | 25.2           | 0.01 | 310 | 5.7                | 0.97 | 53 | 35.7               | 0.85 | 157 |
| Monday                   | 17.8            | –    | 360 | 20.4           | –    | 304 | 3.6                | –    | 56 | 35.0               | –    | 157 |
| Friday                   | 20.2            | 0.40 | 362 | 22.3           | 0.57 | 310 | 7.7                | 0.34 | 52 | 35.4               | 0.93 | 158 |

% = proportion of households fully responding by web. *P* from Chi<sup>2</sup> tests adjusted for sample design.

completed: 56.6% vs. 64.2%,  $P = 0.03$ ), and more refusals (12.4% vs. 6.9%,  $P = 0.03$ ). In contrast for the refreshment sample, partially-responding households were less likely with the mixed mode design (15.6% vs. 25.0%,  $P = 0.01$ ).

There is a suggestion that higher incentive levels might counter-balance the negative effect of the mixed mode design on complete household response (Table 7): with a £10 incentive, the complete household response rate is 1.5 percentage points lower with the mixed mode design, while with a £30 incentive it is 10.3 percentage points higher. However, these differences do not reach statistical significance due to the small sample sizes of households within incentive groups in the refreshment sample.

#### 4.2 Individual Response Rates

Table 6 shows response rates to the individual interview for persons aged 16 or over. Amongst previous wave respondents, the mixed mode design resulted in a higher proportion of refusals than the single-mode face-to-face design (12.5% vs. 6.3%,  $P = 0.03$ ) and amongst previous wave non-respondents it resulted in a smaller proportion of proxy interviews (9.4% vs. 20.2%,  $P = 0.00$ ). The proportion of interviews of any form (full, proxy or partial) was lower with mixed modes for the original sample (65.2% vs. 73.0%,  $P = 0.01$ ), but there was no difference for the refreshment sample. In the refreshment sample there is a suggestion that with higher-value incentives, response rates with mixed modes might be higher than face-to-face, but with lower value incentives, lower than face-to-face. These differences are, however, not statistically significant (Table 7).

The overall pattern regarding response rates to the individual interview therefore seems to be that, compared to face-to-face single-mode, this mixed mode design produces lower response rates for previous wave respondents, though there is no evidence of a difference in response rates either for previous wave non-respondents or for the refreshment sample.

Response to the youth questionnaire for 10 - 15 year-olds

was significantly lower with the mixed mode design compared to face-to-face (33.2% vs. 63.0%,  $P = 0.00$ ). In the mixed mode sample, very few young people responded by web: of all youth respondents in the mixed mode sample, 11% responded by web and 89% completed a paper questionnaire. The youth response rate did not differ between households in which all adults completed the web questionnaire and other households in the mixed mode sample, but the proportion who responded online did differ: in households where all adults participated by web, 8 out of 33 (24%) completed youth questionnaires were completed online, whereas in other mixed mode households none of the 39 youth questionnaires were completed online.

#### 4.3 Response Rates for Subgroups

Though there is no evidence that the mixed mode design increased response rates overall, there could be a positive effect for subgroups. To investigate this we developed multinomial logistic regression models in which we tested the effects of various household characteristics and the interactions of those characteristics with treatment. The dependent variable distinguished between four outcomes: complete household response (household interview and all adult interviews – regardless of whether or not any 10-15 year-olds in the household completed the youth questionnaire), partial household response (at least one adult interview missing), refusal, other non-response. Results for the original sample are summarised in Table 8. With mixed modes, single-parent households and households with children but no couple or related adults were less likely to be fully responding and more likely to be partially responding (despite, recall, only adult response being considered in categorising a household as fully or partially responding). The same was true for households in which all adults were regular web users. For none of the household characteristics analysed did the mixed mode design affect the refusal rate.

For individual-level response we fitted a logit model predicting a full response (versus partial, proxy or non-response)

Table 5  
*Household Response Rates*

|                    | Original sample |      |      |                |      |      |                    |      |      |                    |      |      |
|--------------------|-----------------|------|------|----------------|------|------|--------------------|------|------|--------------------|------|------|
|                    | Total           |      |      | IP4 responding |      |      | IP4 non-responding |      |      | Refreshment sample |      |      |
|                    | F2F             | MM   | P    | F2F            | MM   | P    | F2F                | MM   | P    | F2F                | MM   | P    |
| HH response rate   | 78.0            | 74.2 | 0.22 | 84.1           | 81.1 | 0.31 | 32.6               | 36.0 | 0.69 | 85.1               | 81.9 | 0.45 |
| Complete HHs       | 58.2            | 51.1 | 0.03 | 64.2           | 56.6 | 0.03 | 14.0               | 20.7 | 0.38 | 60.1               | 66.3 | 0.23 |
| Partial HHs        | 19.8            | 23.1 | 0.20 | 19.9           | 24.6 | 0.10 | 18.6               | 15.3 | 0.60 | 25.0               | 15.6 | 0.01 |
| Non-contact        | 8.0             | 8.4  | 0.80 | 6.5            | 5.5  | 0.50 | 18.6               | 24.3 | 0.46 | 6.5                | 7.6  | 0.70 |
| Refusal            | 11.3            | 15.8 | 0.08 | 6.9            | 12.4 | 0.03 | 44.2               | 35.1 | 0.25 | 7.1                | 8.3  | 0.68 |
| Other non-response | 2.7             | 1.5  | 0.24 | 2.5            | 1.0  | 0.07 | 4.7                | 4.5  | 0.97 | 1.2                | 2.2  | 0.43 |
| N                  | 364             | 726  |      | 321            | 615  |      | 43                 | 111  |      | 168                | 315  |      |

F2F = face-to-face; MM = mixed modes; HH = household; P = *P*-values from Chi<sup>2</sup> tests adjusted for sample design.

Table 6  
*Individual Response Rates (including eligible adults in non-respondent households)*

|                    | Original sample |      |      |                |      |      |                    |      |      |                    |      |      |
|--------------------|-----------------|------|------|----------------|------|------|--------------------|------|------|--------------------|------|------|
|                    | Total           |      |      | IP4 responding |      |      | IP4 non-responding |      |      | Refreshment sample |      |      |
|                    | F2F             | MM   | P    | F2F            | MM   | P    | F2F                | MM   | P    | F2F                | MM   | P    |
| Full interview     | 65.6            | 60.6 | 0.12 | 82.8           | 77.1 | 0.10 | 18.5               | 24.7 | 0.13 | 71.3               | 74.4 | 0.46 |
| Proxy interview    | 7.4             | 4.6  | 0.00 | 2.6            | 2.0  | 0.50 | 20.2               | 9.4  | 0.00 | 5.6                | 1.9  | 0.00 |
| Full or proxy      | 73.0            | 65.2 | 0.01 | 85.4           | 79.1 | 0.06 | 38.8               | 34.1 | 0.31 | 76.8               | 76.3 | 0.90 |
| Non-contact        | 8.1             | 6.9  | 0.48 | 5.7            | 3.5  | 0.11 | 15.7               | 15.6 | 0.97 | 4.7                | 6.8  | 0.41 |
| Refusal            | 14.6            | 21.3 | 0.01 | 6.3            | 12.5 | 0.03 | 38.2               | 41.2 | 0.51 | 11.1               | 9.9  | 0.68 |
| Other non-response | 4.3             | 6.6  | 0.09 | 2.6            | 4.8  | 0.15 | 7.3                | 9.1  | 0.47 | 7.3                | 7.0  | 0.86 |
| N                  | 700             | 1412 |      | 494            | 930  |      | 178                | 405  |      | 341                | 587  |      |

F2F = face-to-face; MM = mixed modes; HH = household; P = *P*-values from Chi<sup>2</sup> tests adjusted for sample design.

using individual characteristics and the interactions of those characteristics with treatment. Results for the original sample are summarised in Table 9. Joint tests suggest that respondent characteristics associated with higher response probabilities in face-to-face were also associated with higher response probabilities in mixed modes, with two exceptions:

- With face-to-face, the predicted response rate was 10 percentage points higher in rural than urban locations; in mixed modes the opposite was the case: the predicted rate was 10 percentage points higher in urban locations;
- Respondents who said at the previous wave that they would definitely not respond to a web survey had the highest predicted response rates in the face-to-face treatment, while respondents who said they definitely would respond to a web survey had the highest rates in mixed modes.

However, we have not been able to identify any sample subgroup that was *more* likely to participate with the mixed mode treatment than with face-to-face single-mode. But several groups were *less* likely to give an interview in the mixed

mode treatment: men (-7.0% points), white (-6.6% points), in rural location (-21.3% points), web users (-7.7% points), those for whom we had an email address (-6.5% points), age 21-30 (-20.9%) single with children (-21.0% points), couples with children (-11.9% points), 2+ unrelated adults with children (-25.6% points), individuals who said they would definitely not do the survey by web (-19.7% points).

## 5 Item Nonresponse

In this section, we compare the face-to-face and mixed mode treatments in terms of two measures of item missing data. The first is an overall item non-response rate. The second is the proportion of employed respondents who did not provide a value of their last gross pay. This is an important item as it is central to the derivation of income measures and is known to be prone to relatively high item non-response rates.

The mean item non-response rate is based on 1055 items in the individual questionnaire, of which 6 items about un-

Table 7  
Refreshment Sample Response Rates by Value of Unconditional Incentive

|                          | £10 incentive |      |      | £20 incentive |      |      | £30 incentive |      |      |
|--------------------------|---------------|------|------|---------------|------|------|---------------|------|------|
|                          | F2F           | MM   | P    | F2F           | MM   | P    | F2F           | MM   | P    |
| HH response rate         | 83.3          | 73.3 | 0.24 | 81.0          | 79.3 | 0.79 | 90.5          | 91.2 | 0.88 |
| Complete HHs             | 54.8          | 53.3 | 0.89 | 58.7          | 67.6 | 0.28 | 65.1          | 75.4 | 0.20 |
| Partial HHs              | 28.6          | 20.0 | 0.34 | 22.2          | 11.7 | 0.11 | 25.4          | 15.8 | 0.18 |
| Non-contact              | 4.8           | 11.1 | 0.24 | 11.1          | 7.2  | 0.38 | 3.2           | 5.3  | 0.52 |
| Refusal                  | 11.9          | 13.3 | 0.83 | 7.9           | 11.7 | 0.47 | 3.2           | 0.9  | 0.27 |
| Other non-response       | 0.0           | 2.2  | 0.34 | 0.0           | 1.8  | 0.28 | 3.2           | 2.6  | 0.83 |
| N                        | 42            | 90   |      | 63            | 111  |      | 63            | 114  |      |
| Individual response rate |               |      |      |               |      |      |               |      |      |
| Full interview           | 66.3          | 62.7 | 0.66 | 69.2          | 74.6 | 0.42 | 76.8          | 82.7 | 0.31 |
| Proxy                    | 5.8           | 2.5  | 0.22 | 6.9           | 0.0  | 0.00 | 4.0           | 3.1  | 0.69 |
| Non-contact              | 3.5           | 10.6 | 0.15 | 7.7           | 6.0  | 0.68 | 2.4           | 4.9  | 0.38 |
| Refusal                  | 15.1          | 16.1 | 0.89 | 14.6          | 13.9 | 0.92 | 4.8           | 1.8  | 0.19 |
| Other non-response       | 9.3           | 8.1  | 0.82 | 1.5           | 5.5  | 0.22 | 12.0          | 7.6  | 0.19 |
| N                        | 86            | 161  |      | 130           | 201  |      | 125           | 225  |      |

F2F = face-to-face; MM = mixed modes; HH = household;  $P$  =  $P$ -values from  $\chi^2$  tests adjusted for sample design.

earned income sources were repeated for each source reported. The rate is the proportion of items for which the respondent was eligible, to which they answered “don’t know” or “refused”<sup>3</sup>. Mean item non-response rates were low but around 65% higher with the mixed-mode treatment than with single-mode face-to-face interviewing (Table 10): 1.90% vs. 1.15% for the original sample ( $P = 0.00$ ) and 1.38% vs. 0.82% for the refreshment sample ( $P = 0.00$ ). The item non-response rate for last gross pay was also significantly higher with mixed modes than face-to-face in both the original sample (17.6% vs. 10.0%;  $P = 0.02$ ) and the refreshment sample (18.0% vs. 8.0%;  $P = 0.03$ ).

## 6 Summary and Discussion

Our findings suggest that a mixed mode design including web, along the lines of the one tested in this study, has potential to deliver real cost savings. However, avoiding damage to long-term participation rates and to item response rates may prove more challenging. We discuss the issues of cost savings, unit nonresponse and item nonresponse in turn.

### 6.1 Complete Household Response by Web

As nearly one in five of the original sample households responded entirely by web, accounting for one in four of all fully-responding households, the potential for cost savings clearly exists. Most of these households did not require any interviewer visit. Moreover, our findings hint that even higher web completion rates may be possible if respondents

are offered higher-value unconditional monetary incentives and that the proportion of households who respond fully by web would further increase if all were offered a conditional incentive for completing online. Additionally, the proportion likely to complete fully by web may increase over time due to increases in the proportions who have broadband access, who are regular web users, and who provide the survey organisation with an email address.

### 6.2 Unit Response Rates

Our mixed mode design resulted in a smaller proportion of households fully responding, a smaller proportion of previously-cooperative individuals responding, and more refusals, relative to the face-to-face single-mode design. We have been unable to identify any sample subgroup for which either the household or individual response rate was higher with the mixed mode design. This does not bode well. However, with the higher incentive levels in the refreshment sample, both individual and household response rates were no lower with mixed modes than with face-to-face interviewing. We also found that additional conditional incentives can increase the proportion of households participating fully by web. The conditional incentive offered in our study was

<sup>3</sup> The base excludes modules asked in the self-completion section of the questionnaire, and modules on childcare, newborns, parental expectations and parenting styles which are asked about each child.

Table 8  
*Predicted Probability of Household Response Outcome – Original Sample*

|   | F2F  | Mixed modes | Difference | P-value | P-value<br>(joint test) |
|---|------|-------------|------------|---------|-------------------------|
| <i>Pr(complete HH)</i>                          |      |             |            |         |                         |
| Rural   | 61.2 | 49.6        | -11.5      | 0.11    |                         |
| Urban   | 59.1 | 53.3        | -5.8       | 0.10    | 0.43                    |
| No HH internet                                  | 60.0 | 55.1        | -5.0       | 0.48    |                         |
| HH internet                                     | 59.6 | 51.2        | -8.3       | 0.07    | 0.70                    |
| No emails known                                 | 58.0 | 51.7        | -6.4       | 0.29    |                         |
| Some emails known                               | 56.6 | 49.7        | -7.0       | 0.25    |                         |
| All emails known                                | 63.9 | 56.9        | -7.1       | 0.24    | 0.99                    |
| Single, no children                             | 76.6 | 76.2        | -0.4       | 0.96    |                         |
| Single, children                                | 78.2 | 38.9        | -39.3      | 0.00    |                         |
| Couple, no children                             | 58.8 | 57.6        | -1.2       | 0.84    |                         |
| Couple, children                                | 47.1 | 36.7        | -10.4      | 0.11    |                         |
| 2+ unrelated adults, no children                | 44.1 | 46.9        | 2.8        | 0.73    |                         |
| 2+ unrelated, children                          | 63.2 | 21.2        | -41.9      | 0.00    | 0.00                    |
| No web users, incl. some unknown <sup>a</sup>   | 47.0 | 40.6        | -6.3       | 0.34    |                         |
| Some web users                                  | 72.4 | 64.3        | -8.1       | 0.38    |                         |
| All web users                                   | 79.9 | 66.8        | -13.1      | 0.01    |                         |
| Some unknown, all others web users <sup>a</sup> | 23.1 | 35.3        | 12.2       | 0.09    | 0.05                    |
| <i>Pr(partial HH)</i>                           |      |             |            |         |                         |
| Rural   | 22.9 | 21.8        | -1.1       | 0.84    |                         |
| Urban   | 18.4 | 23.3        | 5.0        | 0.07    | 0.30                    |
| No HH internet                                  | 11.4 | 17.2        | 5.8        | 0.33    |                         |
| HH internet                                     | 21.2 | 24.0        | 2.8        | 0.36    | 0.69                    |
| No emails known                                 | 12.9 | 19.7        | 6.8        | 0.11    |                         |
| Some emails known                               | 28.7 | 30.9        | 2.3        | 0.64    |                         |
| All emails known                                | 13.3 | 16.4        | 3.1        | 0.40    | 0.75                    |
| Single (with and without children)              | 3.1  | 5.3         | 2.2        | 0.57    |                         |
| Couple, no children                             | 22.8 | 26.6        | 3.8        | 0.49    |                         |
| Couple, children                                | 23.8 | 31.0        | 7.2        | 0.24    |                         |
| 2+ unrelated adults, no children                | 26.9 | 26.7        | -0.3       | 0.97    |                         |
| 2+ unrelated, children                          | 14.7 | 37.9        | 23.2       | 0.00    | 0.14                    |
| No web users, incl. some unknown <sup>a</sup>   | 31.0 | 29.4        | -1.6       | 0.81    |                         |
| Some web users                                  | 9.8  | 12.9        | 3.1        | 0.50    |                         |
| All web users                                   | 6.5  | 18.1        | 11.6       | 0.00    |                         |
| Some unknown, all others web users <sup>a</sup> | 31.4 | 29.4        | -2.0       | 0.73    | 0.12                    |
| <i>Pr(refusal HH)</i>                           |      |             |            |         |                         |
| Rural   | 11.0 | 20.9        | 9.9        | 0.09    |                         |
| Urban   | 10.8 | 13.5        | 2.8        | 0.34    | 0.28                    |
| No HH internet                                  | 12.2 | 11.3        | -0.9       | 0.89    |                         |
| HH internet                                     | 10.5 | 17.0        | 6.6        | 0.09    | 0.39                    |
| No emails known                                 | 10.3 | 18.5        | 8.2        | 0.11    |                         |
| Some emails known                               | 8.3  | 14.8        | 6.5        | 0.17    |                         |
| All emails known                                | 15.5 | 12.2        | -3.3       | 0.58    | 0.41                    |
| Single, no children                             | 10.8 | 15.3        | 4.5        | 0.51    |                         |
| Single, children                                | 11.7 | 26.2        | 14.5       | 0.20    |                         |
| Couple, no children                             | 8.4  | 12.6        | 4.1        | 0.26    |                         |
| Couple, children                                | 16.7 | 17.3        | 0.7        | 0.92    |                         |
| 2+ unrelated adults, no children                | 14.3 | 13.0        | -1.3       | 0.80    |                         |
| 2+ unrelated, children                          | 8.1  | 21.8        | 13.6       | 0.10    | 0.47                    |
| No web users, incl. some unknown <sup>a</sup>   | 15.4 | 19.7        | 4.4        | 0.46    |                         |
| Some web users                                  | 10.1 | 14.0        | 3.9        | 0.54    |                         |
| All web users                                   | 5.4  | 9.7         | 4.3        | 0.21    |                         |
| Some unknown, all others web users <sup>a</sup> | 16.7 | 18.7        | 2.0        | 0.75    | 0.99                    |

Estimates from a multinomial logit model, including treatment, household characteristics and interactions between characteristics and treatment as predictors. Probabilities predicted using the "margins" command in Stata version 12. *P*-values from Wald tests adjusted for sample design. *n* = 1,049 households (352 F2F and 697 mixed mode).

<sup>a</sup> "unknown" refers to individuals within the household for whom the web use variable is missing due to item or unit non-response.

Table 9  
*Predicted Response Probabilities – Original Sample, IP4 respondents (%)*

|                        | F2F  | Mixed modes | Difference | P-value | P-value<br>(joint test) |
|------------------------|------|-------------|------------|---------|-------------------------|
| Female                 | 82.8 | 77.8        | -5.0       | 0.22    |                         |
| Male                   | 83.3 | 76.3        | -7.0       | 0.05    | 0.61                    |
| Non-white              | 63.9 | 70.6        | 6.7        | 0.54    |                         |
| White                  | 84.3 | 77.7        | -6.6       | 0.05    | 0.23                    |
| Not in work            | 85.4 | 75.4        | -10.0      | 0.04    |                         |
| In work                | 81.5 | 78.5        | -3.0       | 0.39    | 0.15                    |
| Rural                  | 91.1 | 69.8        | -21.3      | 0.00    |                         |
| Urban                  | 81.1 | 79.3        | -1.8       | 0.61    | 0.00                    |
| Not web user           | 76.2 | 76.5        | 0.3        | 0.96    |                         |
| Web user               | 85.1 | 77.4        | -7.7       | 0.03    | 0.26                    |
| No email given         | 77.2 | 74.1        | -3.1       | 0.59    |                         |
| Email given            | 85.6 | 79.2        | -6.5       | 0.07    | 0.59                    |
| Age 16-20              | 57.5 | 72.5        | 15.0       | 0.32    |                         |
| Age 21-30              | 77.5 | 56.6        | -20.9      | 0.02    |                         |
| Age 31-40              | 78.1 | 72.8        | -5.3       | 0.55    |                         |
| Age 41-50              | 83.1 | 77.3        | -5.9       | 0.28    |                         |
| Age 51-60              | 88.3 | 80.7        | -7.7       | 0.13    |                         |
| Age 61-70              | 90.6 | 86.7        | -3.9       | 0.44    |                         |
| Age 71+                | 86.1 | 81.2        | -4.9       | 0.55    | 0.63                    |
| Single                 | 79.8 | 81.6        | 1.8        | 0.77    |                         |
| Single, children       | 89.9 | 68.9        | -21.0      | 0.02    |                         |
| Couple                 | 84.6 | 85.0        | 0.4        | 0.94    |                         |
| Couple, children       | 82.9 | 71.1        | -11.9      | 0.04    |                         |
| 2+ unrelated adults    | 79.1 | 79.2        | 0.1        | 0.99    |                         |
| 2+ unrelated, children | 84.6 | 59.0        | -25.6      | 0.01    | 0.12                    |
| Web: no                | 91.3 | 71.6        | -19.7      | 0.00    |                         |
| Web: maybe             | 76.1 | 79.3        | 3.2        | 0.50    |                         |
| Web: yes               | 79.5 | 79.7        | 0.2        | 0.98    | 0.00                    |

*N*=1408. Predicted probability of giving a full interview, based on a logit model including the allocated mode, characteristics of the sample members, and interactions between the mode and characteristics as predictors. Predicted probabilities calculated using the command `-margins-` in Stata (version 12). *P*-values from Wald tests adjusted for sample design.

Table 10  
*Item Non-Response Rates by Sample and Treatment*

|                                 | Original sample |      |      | Refreshment sample |      |      |
|---------------------------------|-----------------|------|------|--------------------|------|------|
|                                 | F2F             | MM   | P    | F2F                | MM   | P    |
| Mean Item non-response (%)      | 1.15            | 1.90 | 0.00 | 0.82               | 1.38 | 0.00 |
| N                               | 459             | 856  |      | 243                | 437  |      |
| Gross pay Item non-response (%) | 10.0            | 17.6 | 0.02 | 8.0                | 18.0 | 0.03 |
| N                               | 229             | 410  |      | 112                | 206  |      |

F2F = face-to-face; MM = mixed modes; P=*P*-values from Wald tests of means adjusted for sample design.

small in value relative to the unconditional incentive that each sample member had already received. Future research could usefully investigate whether a larger conditional incentive, representing a reallocation of survey resources from the unconditional to the conditional incentive, might better achieve the objectives of a mixed-mode survey. More generally, there appears to be scope for further refining the incentive offer in order to deliver an appropriate trade-off between costs and nonresponse. Additionally, we found suggestive evidence that the web mode may appeal to different sample members than the CAPI mode. Despite fewer than one in four sample members responding online, those who had been hard-to-get face-to-face at the previous wave were just as likely to respond online as those who had been easy-to-get. This is promising as it suggests that web does not just get responses from those who would in any case have been easy to get face-to-face.

A successful mixed-mode design is likely to involve restricting the request for web participation to certain sample subgroups (Lynn, 2015). These subgroups should be ones that have a relatively high propensity to respond by the web (to achieve cost savings) but for whom a web invite does not reduce their overall propensity to participate in the survey (in order to avoid raising overall attrition rates compared to face-to-face interviewing). Identifying such subgroups remains a challenge as our findings suggest that subgroups that meet one of these criteria may not meet the other. For example, in section 3 we found that households with internet access at home, and whose members are internet users and have provided email addresses had a higher propensity to respond by web, but in section 4.3 we found that those same characteristics were associated with a lower overall response propensity with the mixed-mode design than with the face-to-face single mode design. The findings of section 4.3 also suggest that subgroups such as single-parent families, 21-30 year-olds and people in rural areas may be amongst the least appropriate to ask to participate online.

The introduction of a telephone approach as an additional stage in the sequential design could also help to boost participation rates. This could be done as a final step after all usual efforts have been made, unsuccessfully, to achieve a face-to-face interview. In the context of a mixed-mode survey in which web and face-to-face are the primary modes, the telephone approach could involve encouraging the sample member to participate by web or offering a telephone interview instead.

In the mixed-mode sample, there was very low response to the youth self-completion questionnaire. Gaining cooperation when contact must be via a gatekeeper – the parent, in this case – is always a challenge, but it may be particularly difficult when the gatekeeper’s mode of response is not known in advance. The approach used in our study was not successful and alternative approaches should be identi-

fied and tested.

### 6.3 Item Nonresponse

We find significantly higher levels of non-substantive responses (“don’t know”s and refusals) with the mixed mode design. For gross pay, the difference in levels was particularly substantial. For both measures, differences were significant for both the original and refreshment samples. These differences are consistent with previous research comparing self-administration with interviewer administration (Bowling, 2005; De Leeuw, 2005). This suggests that the longitudinal context may do nothing to reduce the problem of item nonresponse in self-administered questionnaires.

Future research could usefully investigate ways of reducing item non-response with web data collection in longitudinal surveys. Our approach of simply having the “Don’t Know” and “Refuse” options appear on a follow-up screen whenever a respondent attempted to skip a question may not be the best way to encourage item response. It could be worthwhile to experiment with alternatives, in the spirit of Oudejans and Christian (2011) and Smyth, Dillman, Christian, and McBride (2009). This might involve testing alternative forms of the message that appears when a skip is attempted, or alternative positioning (e.g. at the end of the questionnaire instead of immediately after the attempt to skip).

### 6.4 Conclusion

This study has demonstrated that transitioning from a face-to-face longitudinal survey to a mixed mode web and face-to-face survey is not straightforward. Cost savings appear possible but avoiding increased attrition and increased item nonresponse is challenging. Targeting the web invitation to a subset of the sample is likely to be more effective than inviting all sample members to the web but, while we have identified some characteristics that could contribute to a targeting strategy, the details of such a strategy remain to be developed. Further experimentation is called for to help to identify the combination of design features and parameters that will constitute a mixed mode approach that can bring cost savings without materially damaging the quality of an ongoing panel survey.

## 7 Acknowledgements

The work reported in this paper is part of the *Understanding Society* programme, funded by the UK Economic and Social Research Council (ESRC). Earlier versions of the paper were presented by Annette Jäckle at the 6th Measurement and Experimentation in the Social Sciences (MESS) workshop, Amsterdam, 31 August – 1 September 2012 and by Peter Lynn at the (UK) Social Research Association annual conference, London, 10 December 2012.

## References

- Betts, P. & Lound, C. (2010). The application of alternative modes of data collection in UK government social surveys: a report for the Government Statistical Service. Newport: Office for National Statistics. Retrieved from <http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/data-collection-methodology/reports-and-publications/alternative-modes-of-data-collection/index.html>
- Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, 27, 281–291.
- Buck, N. & McFall, S. (2012). Understanding Society: design overview. *Longitudinal and Life Course Studies*, 3, 5–17.
- Cantor, D., O'Hare, B., & O'Connor, K. (2008). The use of monetary incentives to reduce non-response in random digit dial telephone surveys. In J. Lepkowski, C. Tucker, J. Brick, E. De Leeuw, L. Japac, P. Lavrakas, ... R. Sangst (Eds.), *Advances in telephone survey methodology* (pp. 471–498). Wiley.
- Church, A. (1993). Estimating the effect of incentives on mail survey response rates: a meta-analysis. *Public Opinion Quarterly*, 57, 62–79.
- Couper, M. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75, 889–908.
- Couper, M. (2012). Assessment of innovations in data collection technology for Understanding Society: a report to the Economic and Social Research Council. Retrieved from <http://eprints.ncrm.ac.uk/2276/>
- Cuppen, M., Van Der Laan, P., & Van Nunspeet, W. (2011). Re-engineering dutch social surveys: from single-purpose surveys to an integrated design. Proceedings of 58th World Statistical Congress, Dublin. The Hague: International Statistical Institute. Retrieved from <http://2011.isiproceedings.org/>
- De Leeuw, E. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21, 233–255.
- Dex, S. & Gurney, J. (2011). *On the experience and evidence about mixing modes of data collection in large-scale surveys where the web is used as one of the modes in data collection*. National Centre for Research Methods Review Paper, University of Southampton.
- Dillman, D. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods*, 17, 30–52.
- Dillman, D. (2009). Some consequences of survey mode changes in longitudinal surveys. In P. Lynn (Ed.), *Methodology of longitudinal surveys* (pp. 127–140). Chichester: Wiley.
- Duffy, B., Smith, K., Terhanian, G., & Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, 47, 615.
- Fricker, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of web and telephone surveys. *Public Opinion Quarterly*, 69, 370–392.
- Görizt, A. (2006). Incentives in web studies: methodological issues and a review. *International Journal of Internet Science*, 1, 58–70.
- Görizt, A. (2010). Using lotteries, loyalty points, and other incentives to increase participant response and completion. In S. Gosling & J. Johnson (Eds.), *Advanced methods for conducting online behavioural research*. Washington DC: American Psychological Association.
- Görizt, A. (2015). Incentive effects. In U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, & P. Sturgis (Eds.), *Improving survey methods: lessons from recent research* (pp. 339–350). London: Routledge.
- Griffin, D., Fischer, D., & Morgan, M. (2001). *Testing an internet response option for the American Community Survey*. Paper presented at the annual conference of the American Association for Public Opinion Research Montreal, Quebec, Canada, May 17–20.
- Groves, R. & Lyberg, L. (2010). Total survey error: past, present, and future. *Public Opinion Quarterly*, 74, 849–879.
- Harris, L., Weinberger, M., & Tierney, W. (1997). Assessing inner-city patients' hospital experiences: a controlled trial of telephone interviews versus mailed surveys. *Medical Care*, 35, 70–76.
- Jäckle, A. & Lynn, P. (2008). Respondent incentives in a multi-mode panel survey: cumulative effects on non-response and bias. *Survey Methodology*, 34, 105–117.
- Kreuter, F. (2013). Facing the nonresponse challenge. *Annals of the American Academy of Political and Social Science*, 645, 23–35.
- Lagerström, B. (2011). *Web in mix mode surveys in Norway*. Paper presented at NCRM/SRN Symposium on General Population Surveys on the Web, London, November 3. Retrieved from <http://eprints.ncrm.ac.uk/1917/>
- Laurie, H. & Lynn, P. (2009). The use of respondent incentives on longitudinal surveys. In P. Lynn (Ed.), *Methodology of longitudinal surveys* (pp. 205–233). Chichester: Wiley.
- Luiten, A. & Blanke, K. (2013). ESSnet project on data collection for social surveys using multi modes (DCSS). Paper presented at Conference on New Techniques and Technologies for Statistics, Brussels, 5–7 March. Retrieved from [http://www.cros-portal.eu/sites/default/files/NTTS2013fullPaper\\_248.pdf](http://www.cros-portal.eu/sites/default/files/NTTS2013fullPaper_248.pdf)
- Lynn, P. (2009). *Sample design for Understanding Society*. Understanding Society Working Paper 2009-01, ISER, University of Essex, Colchester. Retrieved from [www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2009-01](http://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2009-01)

- Lynn, P. (2013). Alternative sequential mixed mode designs: effects on attrition rates, attrition bias and costs. *Journal of Survey Statistics and Methodology*, 1, 183–205.
- Lynn, P. (2015). Targeted response inducement strategies on longitudinal surveys. In U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, & P. Sturgis (Eds.), *Improving survey methods: lessons from recent research* (pp. 322–338). London: Routledge.
- Lynn, P. & Lievesley, D. (1991). *Drawing general population samples in Great Britain*. London: SCPR.
- Lynn, P., Uhrig, S., & Burton, J. (2010). *Lessons from a randomised experiment with mixed mode designs for a household panel survey*. Understanding Society Working Paper 2010-03, ISER, University of Essex, Colchester. Retrieved from [www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2010-03](http://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2010-03)
- Nicolaas, G., Thomson, K., & Lynn, P. (2000). *The feasibility of conducting electoral surveys in the UK by telephone*. London: Centre for Research into Elections and Social Trends.
- Oudejans, M. & Christian, L. (2011). Using interactive features to motivate and probe responses to open-ended questions. In M. Das, P. Ester, & L. Kaczmirek (Eds.), *Social and behavioral research and the internet* (pp. 215–244). Hove: Routledge.
- Schoeni, R., Stafford, F., McGonagle, K., & Andreski, P. (2013). Response rates in national panel surveys. *Annals of the American Academy of Political and Social Science*, 645, 60–87.
- Singer, E., Gebler, N., Raghunathan, T., Van Hoewyk, J., & McGonagle, K. (1999). The effect of incentives in interviewer-mediated surveys. *Journal of Official Statistics*, 15, 217–230.
- Singer, E. & Ye, C. (2013). The use and effects of incentives in surveys. *Annals of the American Academy of Political and Social Science*, 645, 112–141.
- Smyth, J. D., Dillman, D. A., Christian, L. M., & McBride, M. (2009). Open-ended questions in web surveys: can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly*, 73, 325–337.
- Tourangeau, R., Conrad, F., & Couper, M. (2013). *The science of web surveys*. Oxford: Oxford University Press.
- Uhrig, S. (2011). Using experiments to guide decision making in Understanding Society: introducing the Innovation Panel. Colchester: University of Essex. Retrieved from <http://www.understandingsociety.ac.uk/research/publications/findings/early>