Understanding Society
Working Paper Series

No. 2022 – 07

September 2022

Understanding Society Wave 14 Boost Trial: Experiments with methods of recruiting a probability online boost sample

Joel Williams[1], Liz Ward[1], Jonathan Burton[2], Hannah Carpenter[1], Kirsty Cole[3], Bruce Hayward[1], Gerry Nicolaas[3], Violetta Parutis[2], Bronwyn Seymour[1], Katie Thornton[1], Martin Wood[3], and Laura Woodward[1]

[1]Kantar Public, [2]University of Essex, [3]NatCen Social Research

Understanding Society

Economic and Social Research Council

Non-technical summary

This paper documents the results of experiments carried out in the Wave 14 (W14) pilot, designed to test different methods of recruiting a boost sample for *Understanding Society* using a web survey design. The W14 web pilot was conducted in October 2021 with an issued sample of 17,304 UK addresses. The pilot contained a number of overlapping randomised controlled trials, each testing one particular treatment designed to raise the response rate.

Overall, there were six treatments designed to encourage *any* response, and another two designed to encourage complete response from households that had supplied at least some data.

Only two of the six treatments designed to encourage any response had an appreciable impact:

- variations in the contact sequence (inclusion of a pre-notification letter and/or an additional (third) reminder);
- the offer of an early bird bonus incentive of £10 over a base £20, instead of a uniform £30;

The other four treatments had little or no impact:

- adding sponsor logo(s) to the envelope and letter (University of Essex and ISER);
- including post-it-notes in the invitation letter (with variations in message and logo);
- revealing the longitudinal nature of the survey in the invitation letter rather than at a later point (IP14 subsample only); and
- revealing the 'all adults' nature of the survey in the invitation letter rather than revealing it to the first respondent once he/she had completed the household grid.

The two treatments designed to encourage complete response also had little or no impact:

- asking the first respondent to provide the email addresses of other adults in the household; and
- including encouraging messages at intervals during the survey.

The results of these experiments informed the choice of protocols that are being used for the *Understanding Society* boost sample, recruited as part of Wave 14 in 2022-3.

Understanding Society Wave 14 Boost Trial: Experiments with methods of recruiting a probability online boost sample

Joel Williams[1], Liz Ward[1], Jonathan Burton[2], Hannah Carpenter[1], Kirsty Cole[3], Bruce Hayward[1], Gerry Nicolaas[3], Violetta Parutis[2], Bronwyn Seymour[1], Katie Thornton[1], Martin Wood[3], and Laura Woodward[1]

[1]Kantar Public, [2]University of Essex, [3]NatCen Social Research

Abstract: This paper summarizes experiments carried out as part of a pilot survey that was designed to pre-test different protocols for recruiting a probability-based sample of households via a web survey. The pilot was conducted in October 2021 with an issued sample of 17,304 UK addresses. The pilot included six experiments testing different fieldwork protocols designed to increase response rates. Only two of the six treatments had an appreciable impact: variations in the contact sequence (inclusion of a pre-notification letter and/or an additional (third) reminder) and the offer of an early bird bonus incentive of £10 over a base £20, instead of a uniform £30.

Keywords: refreshment sample, push-to-web survey, fieldwork protocols, longitudinal survey.

JEL classification: C80, C81, C83

Acknowledgements: *Understanding Society* is an initiative funded by the Economic and Social Research Council (ESRC) and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. The wave 14 pilot was funded by ESRC grant ES/T002611/1.

Contact: Jonathan Burton, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK, jburton@essex.ac.uk.

1. Introduction

Understanding Society: The UK Household Longitudinal Study is a longitudinal probability sample survey that interviews all adult members of sample households annually about their health, education, work, income, family, housing and civic engagement. The data from the *Understanding Society* study are available from the UK Data Service.[1]

The main survey began in 2009 and includes a large General Population Sample plus three other components: the Ethnic Minority Boost Sample, the former British Household Panel Survey sample and, from Wave 6 (2015), the Immigrant and Ethnic Minority Boost Sample. *Understanding Society* also includes a separate Innovation Panel: a sample of 1,500 households in Great Britain, that is used for methodological testing and experimentation. In wave 1 the *Understanding Society* samples included 40,000 interviewed households. Over time the number of respondent households has fallen due to attrition, as well as sample members becoming ineligible (deceased or moving abroad), although children born to sample members and people moving into sample households also become part of the sample. To replenish the main *Understanding Society* sample a boost (refreshment) sample is being planned for wave 14 (2022-3).

This paper reports on a series of experiments that were carried out in a pilot study in September 2021. The experiments tested different protocols for recruiting a probability sample for a web survey. The results informed the choice of protocols that are being used for the 2022-3 *Understanding Society* boost sample.

2. The Wave 14 *Understanding Society* web trial

Sample

The 2021 web trial was based on 17,304 issued UK addresses, stratified by local measures of deprivation and ethnic diversity. The sample included two sub-samples: 6,047 of these addresses were selected as a clustered sample that will be used as a refreshment sample for the Innovation Panel (IP14 refreshment sample); the other 11,257 addresses formed an additional (unclustered) sample that will not be followed-up in future waves of the survey.

Experiments

The wave 14 (W14) web trial contained **eight overlapping randomised controlled trials**, each testing one particular treatment designed to raise the response rate. Six of the eight were designed to encourage households to at least get started (to provide *any* data), while two of the eight were designed to encourage already participating households to complete their data submission (i.e., for all adults to complete the individual module).

[1] https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000053

The six treatments designed to encourage any response were:

- variations in the contact sequence (inclusion of a pre-notification letter and/or an additional (third) reminder);
- the offer of an early bird bonus incentive of £10 over a base £20, instead of a uniform £30;
- adding sponsor logo(s) to the envelope and letter (University of Essex and ISER);
- including post-it-notes in the invitation letter (with variations in message and logo);
- (IP14 only) revealing the longitudinal nature of the survey in the invitation letter rather than at a later point; and
- revealing the 'all adults' nature of the survey in the invitation letter rather than revealing it to the first respondent once he/she had completed the household grid.

The two treatments designed to encourage complete response were:

- asking the first respondent to provide the email addresses of other adults in the household (although, in the trial, these emails were not actually used to contact anybody); and
- including 'encouraging messages' at intervals during the survey.

Details of the design and implementation of each experiment are documented along with the results in Section 3.

Analysis methods

All results are design-weighted to better reflect the equal probability sample design for the W14 boost survey. All quoted confidence intervals account for this weighting as well as the sample stratification and clustering. All the conversion and response rates quoted in this paper are averaged across the trials.

3. Results

Overall results averaged across all experimental treatments

In total, **12.4% of addresses yielded at least a household grid** (95CI = 11.7%-13.2%). Given that c.9% of sampled addresses will not have contained a residential household (i.e., they are 'deadwood'), this address conversion rate can be converted into an 'any data' household response rate of 13.7%.[2]

The vast majority (87%) of addresses yielding a household grid also yielded full data about the household as well as individual level data from at least one adult. *Complete* data collection (grid + household data + data from *all* adult individuals) was rarer because, in most households, more than one adult needed to complete the individual level module. The

[2] "Deadwood" addresses are those at which there is no residential household. They include addresses where there is no house, or a house under construction; a demolished/derelict house; a vacant or empty housing unit; a non-residential address such as a school or business; a holiday/weekend home; or an institutional communal establishment.

probability of complete data collection, given a household grid, was only 54%. Consequently, the complete data response rate was much lower than the 'any data' response rate: 7.4% (95CI = 6.9%-8.0%).

For the rest of this paper, we focus on address conversion rates rather than response rates, given that the denominator of the response rate can only be estimated at the web stage. The **complete data address conversion rate was 6.8%** (95CI = 6.3%-7.3%).

Given a 27-28 address assignment for the W14 boost survey, an interviewer can be expected to get c.24 addresses where all data collection is required, plus one or two addresses where partial data collection is required. On average, two addresses will be removed from the assignment altogether because all data has been collected at the web stage.

Before moving on, it is important to remember that the conversion rates and response rates quoted in this paper are averaged across the (other) trials embedded in the study. Only the most effective treatments will go forward to the W14 boost survey itself so we might expect the web conversion rate to be higher than it was in this trial.

Results in different sample strata

The **sample was stratified by measures of deprivation and ethnic diversity**. Approximately one fifth (21%) of UK addresses are in deprived areas and just under one third (32%) are in (relatively) ethnically diverse areas.[3] Diversity is correlated with deprivation (see table 1).

Table 1: Lower layer super output area (LSOA) ethnic diversity and deprivation strata

Stratum	Share of UK addresses	Share of UK addresses in this diversity stratum	% deprived within this diversity stratum
Diverse & deprived	9.1%	31.5%	29%
Diverse but not deprived	22.4%		
Not diverse but deprived	12.1%	68.5%	18%
Not diverse nor deprived	56.5%		

Overall, the address **conversion rate in deprived areas was significantly lower than in other areas** (9.2% any data, compared to 13.3% for other areas; 5.5% complete data, compared to 7.1% for other areas; null hypothesis p for both measures <0.001).

[3] 'Deprived' = local super output area's index of multiple deprivation is equal to or above the top quintile for that country; 'ethnically diverse' is a Kantar Public-constructed variable: ethnically diverse census output areas are those that, in combination, covered >= 80% of five populations (Indian, Pakistani, Bangladeshi, Black, and Other non-White/Mixed) *and* were <90% White in the 2011 Census.

The **ethnic diversity of the area was less obviously a factor**: the 'any data' address conversion rate was 11.4% in ethnically diverse areas compared to 12.9% in less diverse areas (null p<.05); however, the difference in the complete data rate was only 0.5 percentage points (6.4% in ethnically diverse areas; 6.9% in less diverse areas (null p>.05).

These top-line figures hide an interesting interaction, albeit an interaction which is not statistically significant. **Within deprived areas, addresses in ethnically diverse neighbourhoods yielded *more* data** than was obtained from addresses in less diverse neighbourhoods (9.8% v 8.7% 'any data' conversion rate, but null p>0.05), but the **reverse was true within less deprived areas** (12.0% 'any data' conversion rate for addresses in ethnically diverse neighbourhoods v 13.8% for addresses in less diverse neighbourhoods; null p<.05).

Table 2: Response in diversity and deprivation strata

Stratum	Share of UK addresses	Any data	Complete data
Diverse & deprived	9.1%	9.8%	5.9%
Diverse but not deprived	22.4%	12.0%	6.6%
Not diverse but deprived	12.1%	8.7%	5.2%
Not diverse nor deprived	56.5%	13.8%	7.3%
All UK	100.0%	12.4%	6.8%

Many other area-level statistics can be attached to the sample frame to help understand patterns in web response; assessing stratum-level patterns is a start but by no means an end to that work. However, the focus of *this* paper is on the various design treatments that were incorporated into the trial, so we leave a discussion of general response patterns to another time.

Individual response within 'any data' households

The **mean number of adults within responding households** (i.e., those providing at least a household grid) **was 1.87**. Just under one third (32%) were single-adult households, just over one half (54%) contained two adults, and one in six (16%) contained at least three adults. This distribution is very similar to the national distribution recorded in the UK Labour Force Survey.[4]

Approximately **two thirds (65%) of adults in responding households completed the individual-level module** (a mean of 1.22). Naturally, this was highest in single-adult

[4] https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyuserguidance

households, where the conditional individual-level completion rate was 87%. It was 61% in households with at least two adults, and 57% in households with at least three adults.[5]

For context, the England-only 2020-21 Department for Digital, Culture, Media & Sport (DCMS) Community Life Survey (CLS) – which had an 'any data' web conversion rate of 23% - recorded a mean number of adults per responding household of 2.06, and an average of 1.49 completed[6] web questionnaires in these households (a conditional response rate of 73%).[7] To compare these figures with the W14 web trial, we need to re-base the latter to include only households with at least one responding adult (not all enumerated households). If we do this, we find that the conditional response rate of the W14 web trial was about the same as for the CLS (76% v 73%). The conditional response rate in households with at least two adults was 71% (cf. 68% for the CLS) and it was 65% in households with at least three adults (higher than the CLS figure of 57%).

For the Community Life Survey, four survey logins were provided in the invitation letter so that up to four adults could complete the web survey *without reference to household enumeration data from the first respondent*. For the W14 web trial, more control was exerted over the number of people completing the individual level module, and the method used (single shared serial number; personalised reminders to enumerated but non-responding individuals) appears to have worked quite well, encouraging the majority of adults in enumerated multi-adult households to respond. The risk of multiple responses from the same person is almost certainly lower than with the simpler CLS method, due to the enumeration stage and the password-protection of individual-level data.

In one (randomly allocated) half of the W14 web trial invitation letters, the requirement for all adults to complete a questionnaire was made explicit; in the other half, this was only made explicit to the first respondent after completing the household grid. Among responding households, this does not appear to have influenced the structure of the household grid: the **mean number of enumerated adults was 1.89 if the 'all-adult' nature of the survey was explicit in the letter, and 1.85 otherwise** (null p>.05). This is a positive finding as there was a risk that revealing the multiple-response/multiple-incentive nature of the survey *before* the enumeration stage might prompt some first respondents to inflate the size of the household when completing the grid, opening up the potential for more incentives.

[5] This is different from the 'complete data' rate measured at a household-level. The conditional individual-level completion rate is the sum of adults completing the individual level module divided by the sum of adults recorded in all household grids.

[6] Completed *and passed a QC check* designed to remove poor quality data as well as suspected additional responses from the same person. Note also that the CLS could be completed on paper, but the results reported here ignore all paper responses.

[7] https://www.gov.uk/government/statistics/community-life-survey-202021

Trial 1: Varying the contact sequence

There is substantial evidence that reminders and/or prenotification letters can be effective at raising response levels to letter-based surveys. However, there is not much evidence about their *relative* impact.

The advantage of a pre-notification letter is that it can focus on the study rather than the details of participation and this may enhance the legitimacy of subsequent invitation letters. In contrast, an invitation or reminder letter has to work as a specific call-to-action, reserving space for elements like login details and incentive claim instructions. The cost of a pre-notification letter or a reminder letter is very similar, so which is most effective at raising response?

Three different contact sequences were trialled: (i) an invitation letter plus up to three household-level reminder letters, approximately one week apart; (ii) the same but with a pre-notification letter a week before the invitation letter; and (iii) as (ii) but with no third reminder. The 'any data' and 'complete data' address conversion rates for each treatment are provided in table 3.

Table 3: Response by contact sequence

Treatment	Any data (95CI)	Complete data (95CI)
(i) Invitation letter + 3 reminders	10.9% (10.0%-11.9%)	5.7% (5.1%-6.4%)
(ii) Prenotification letter + invitation letter + 3 reminders	14.1% (12.9%-15.3%)	7.8% (7.0%-8.7%)
(iii) Prenotification letter + invitation letter + 2 reminders	12.4% (11.3%-13.6%)	6.8% (6.1%-7.5%)

From this table, we can see that the addition of a prenotification letter (treatment (ii) versus treatment (i)) raised the 'any data' conversion rate by 3.2 percentage points (95CI = 1.6-4.7 %pts) while the addition of a third reminder (treatment (ii) versus treatment (iii)) raised it by 1.7 percentage points (95CI = 0.1-3.4 %pts). From this, we can see that a prenotification letter has more value than a third reminder, raising the 'any data' conversion rate by an additional 1.5 percentage points relative to a third reminder (95CI = -0.1-3.0 %pts).

A similar story emerges if we focus on the *complete* data rate. The addition of a prenotification letter raised the complete data conversion rate by 2.1 percentage points (95CI = 1.0-3.2 %pts) while the addition of a third reminder raised it by 1.0 percentage points (95CI = -0.1-2.2 %pts). As with the 'any data' rate, the prenotification letter has more

value than a third reminder, raising the 'complete data' conversion rate by an additional 1.1 percentage points relative to a third reminder (95CI = 0.0-2.1 %pts).

Overall, it is clearly worthwhile to send a prenotification letter. The value of a third reminder is much slighter, but the evidence still points to a small positive effect. It should be included in the contact sequence if the budget and schedule allow.

Trial 2: The offer of an early bird incentive

A deadline for response is usually included in survey invitation letters but if this is some weeks away, sampled individuals may 'put off' responding and then never do it.[8] To counter this effect without changing the overall deadline, an 'early bird' bonus may be offered to encourage immediate rather than delayed response.

Two different incentive packages were trialled: (i) a £30 per-person conditional incentive with a five-week deadline (the end of the web-only fieldwork phase), and (ii) a £20 conditional incentive with a five-week deadline but with a £10 'early-bird' bonus if the survey was done within two weeks. The early bird bonus was not mentioned in either of the second or third reminders (where these were used) as the two-week deadline had passed. The 'any data' and 'complete data' address conversion rates for each treatment are provided in table 4.

Table 4: Response by incentive offer

Treatment	Any data (95CI)	Complete data (95CI)
(i) £30 conditional incentive with 5-week deadline	12.7% (11.7%-13.7%)	7.1% (6.4%-7.8%)
(ii) £20 conditional incentive with 5-week deadline plus extra £10 if respond within 2 weeks	12.2% (11.3%-13.2%)	6.4% (5.8%-7.1%)

From this table, we can see that offering a smaller general incentive but with a £10 'early bird' bonus (treatment (ii) versus treatment (i)) reduced the 'any data' conversion rate by 0.5 percentage points and the 'complete data' conversion rate by 0.7 percentage points. Neither of these differences are large enough to reject the null hypothesis (p>.05; 95CI for difference in complete data rates = -1.6 to +0.3 %pts) but the balance of this evidence favours the larger general incentive of £30.

However, the number of completed individual level modules was about the same under both treatments: the mean number per responding household (i.e., those households that

[8] Examples from behavioural science show earlier and more salient deadlines can help boost response for postal voting https://www.ideas42.org/tool/boosting-ballot-returns/

provided at least an enumeration grid) was 1.25 under the early bird £20+£10 treatment and 1.19 under the standard £30 treatment and this was enough to make up for the slightly lower 'any data' rate.

Where an early bird incentive had been offered, two thirds (66%) of the completed individual level modules were done in the first 14 days, compared to only 42% where the standard incentive was offered. The early bird incentive clearly raised the response level before its deadline, even if it did not raise the overall response level. Based on this data, the mean incentive paid out per responding individual was £26.63, a saving of £3.37 compared to the standard incentive. Per responding household, the mean incentive pay-out was £33.31, a saving of £2.27 compared to the standard incentive.

As the W14 boost survey is currently designed, the 'early bird' bonus would be for completing a questionnaire during the web-only phase of fieldwork (first five weeks): i.e., before the household is passed to an interviewer. Although this trial had a quite different design, it suggests that people *may* respond to an incentivised earlier deadline without this causing serious damage to the overall response rate. However, we cannot use this evidence to confidently predict that an early bird bonus that covers the *whole* of the W14 web-only fieldwork phase would lead to a higher web response rate than an alternative in which an incentive of the same value is used but not presented as a bonus and without a deadline. Even if that seems plausible – likely even - we cannot *quantify* the effect.

However, ISER could instead adopt the trial design for the W14 boost survey such that the early bird bonus is just for completing the survey in the first two weeks of the web-only fieldwork period, rather than the whole five weeks. Based on the evidence above, that would compress two thirds of the web responses into the first two weeks and (potentially) allow face-to-face interview fieldwork to begin earlier (e.g. after week four rather than week five). The question is whether there is any real value in adding a week to the interview fieldwork period at the expense of the web fieldwork period. Week five of the web fieldwork period is at least as productive as the final week of the interview fieldwork period and with much lower costs per completed questionnaire.

In summary, it probably makes sense to offer an early bird bonus that covers the whole of the W14 web fieldwork phase but the evidence from this trial is too specific for us to estimate the effect that would have on the W14 web response rate.

Trial 3: Varying the logo on the envelope and letter

A plain white envelope does not look important, special, or interesting. This is less of an issue for the longitudinal panel since the envelope includes the recipient's name, but for the refreshment sample where we do not have the names of sample members, there is a risk that people discard the letter without opening it, or treat it as junk mail. There is evidence that the addition of a logo may have a small positive effect on responses probability, but

also that removal of one logo from a pair can have a positive effect, suggesting that the identity of the sponsor is important.[9]

Three different logo strategies were trialled: (i) no logo; (ii) the University of Essex logo; and (iii) the University of Essex logo combined with an Institute for Social and Economic Research (ISER) logo. These logos were placed on both the envelope and letterheads. The 'any data' and 'complete data' address conversion rates for each treatment are shown in table 5.

Table 5: Response by logo strategy

Treatment	Any data (95CI)	Complete data (95CI)
(i) No logo	12.4% (11.3%-13.5%)	7.1% (6.4%-7.9%)
(ii) University of Essex logo	11.9% (10.9%-13.1%)	6.8% (6.1%-7.6%)
(iii) University of Essex & ISER logos	13.0% (12.0%-14.2%)	6.4% (5.6%-7.2%)

The results are mixed – using both logos yielded the highest 'any data' rate but the lowest 'complete data' rate – but we cannot reject the null hypothesis that logo strategy is irrelevant to response maximisation. Treatment (iii) – both logos - is clearly the most informative approach and might be preferred on that basis.

Trial 4: The inclusion of post-it-notes in the invitation letter (with variations in message and logo)

A physical – and therefore 'bulky' – non-monetary incentive could encourage respondents to open the envelope simply by eliciting curiosity and there is some evidence that people perceive heavier or bulkier postal items to have higher value.[10]

Like monetary incentives, gifts could encourage participation by eliciting a reciprocity effect. A gift that is aligned with the ethos of the survey and includes the *Understanding Society* logo could act as a reminder about the survey, especially if it creates an implementation intention to complete the survey, for example, branded post-it notes that include instructions about when to take part or how.

[9] On the 2015-16 Active Lives Survey, Ipsos added the logos of the sponsors, Sports England and Public Health England, to the envelope partway through fieldwork, non-experimentally. The response increased from 15% to 20%, with no other changes to the protocol during that period. http://doc.ukdataservice.ac.uk/doc/8223/mrdoc/pdf/8223_technical_report_active_lives_survey_year_1.pdfb Kantar tested the use of the Financial Conduct Authority logo on the 2017 Financial Lives Survey, and achieved a 6.5% web response rate, compared to 5.0% for an envelope without the logo. https://www.fca.org.uk/publication/research/financial-lives-survey-2017-technical-report.pdf

[10] See Test 2 of the Labour Market Survey (2018) https://www.ipsos.com/sites/default/files/ct/news/documents/2018-04/ons_test_2_full_report_final_public_170418.pdf

We trialled the inclusion of a small pack of post-it notes with the *Understanding Society* logo - either with or without an encouraging message ('Help shape the decisions that affect people like you… take part in the study today.") – on the basis that it might work as a visual prompt to action, both to open the bulkier letter and then to complete the survey.

Three different post-it strategies were trialled: (i) no post-it notes; (ii) post-it notes with the *Understanding Society* logo; and (iii) post-it notes with the *Understanding Society* logo and an encouraging message. The 'any data' and 'complete data' address conversion rates for each treatment are shown in table 6.

Table 6: Response by post-it notes strategy

Treatment	Any data (95CI)	Complete data (95CI)
(i) No post-it notes	12.4% (11.3%-13.5%)	6.5% (5.8%-7.3%)
(ii) Post-it notes with logo	12.2% (10.9%-13.5%)	6.6% (5.8%-7.5%)
(iii) Post-it notes with logo and encouraging message	12.8% (11.8%-13.9%)	7.1% (6.4%-8.0%)

There is perhaps an indication that putting an encouraging message on the post-it notes pack was helpful but the overall evidence in favour of post-it notes is weak, and they are probably not worth the additional cost.[11]

Trial 5: Revealing the longitudinal nature of the survey in the IP14 invitation letter rather than at a later point

Currently, participants are not advised of the longitudinal nature of the study in the invitation letter and there is mixed evidence about how best to approach this: temporal discounting theory suggests respondents may discount the cost of taking part in future surveys; other evidence suggests that it is better to seek a one-time response and ask participants to commit to further involvement later;[12] some interviewers saw the longitudinal nature of the survey as a selling point.

This trial was limited to the 6,047 issued addresses that were to be used as a refreshment sample for the Innovation Panel, because responding households at these addresses will be invited to participate in the annual Innovation Panel study. The other sample addresses will not be recontacted for future interviews.

Two treatments were trialled: (i) introducing the longitudinal nature of the survey in the information leaflet sent with the invitation letter, and (ii) introducing it only at the end of

[11] At the time of the experiment in 2021, the cost of the post-it notes, if bought in bulk, was around 40 pence per pack with just a logo, and around 70 pence with a logo and message. The post-it notes were thin enough (30 sheets per pack) that they did not add to postage costs, despite adding some bulk to the envelope.

[12] *Thaler,R & Benartzi, S.(2004). Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving Journal of Political Economy, 2004, vol. 112, no. 1, pt. 2.*

the first respondent's survey. For group (i) the information leaflet contained the wording *"Your information is **most valuable** when we learn how it changes over time. We will invite you to complete the study again next year and **hope you will agree to take part**."* For group (ii) the information leaflet did not include these sentences but was otherwise identical. The 'any data' and 'complete data' address conversion rates for each treatment are shown in table 7.

Table 7: Response by longitudinal introduction strategy

Treatment	Any data (95CI)	Complete data (95CI)
(i) Letter introduction	13.5% (11.8%-15.4%)	7.0% (6.0%-8.2%)
(ii) End-of-survey introduction	14.1% (12.4%-16.0%)	7.7% (6.6%-9.0%)
All IP14	13.8% (12.3%-15.4%)	7.4% (6.5%-8.4%)

First, it is worth noting that the IP14 conversion rates were slightly higher than the web trial conversion rates as a whole. The 'any data' conversion rate was 13.8% (cf.12.4% for the web trial as a whole) and the 'complete data' conversion rate was 7.4% (cf.6.8% for the web trial as a whole). These differences are large enough to reject the null hypothesis at the conventional levels ($p<.05$ for the 'complete data' rate, and $p<.01$ for the 'any data' rate). Given that we have applied design weights before analysing this data, it suggests that there was a positive treatment effect for the IP14 against the standard web trial. The only clear distinction between the two was the longitudinal nature of the IP14 study, revealed by letter or to the first respondent. Is it possible that this is a net positive feature, contrary to prior beliefs? If so, the effect is probably small: the complete data rate of 7.0% when the longitudinal nature of the survey is revealed in the letter is only a little higher than the complete data rate of the non-IP web trial (6.3%).

As for the trial itself, there is a *suggestion* that an end-of-survey introduction of the longitudinal nature of Understanding Society is preferable to including it in the invitation letter. However, the statistical evidence is not strong enough to reject the null hypothesis of no-effect and it could be argued that, because a letter-based introduction is more *transparent*, that is the more ethical strategy and that it should be taken.

Trial 6: Revealing the 'all adults' nature of the survey in the invitation letter rather than revealing it to the first respondent once he/she had completed the household grid

Understanding Society requires all resident adults to complete the individual-level module before household data collection can be considered 'complete'. However, there is scope to vary the point at which this feature is introduced.

Different approaches have been tested regarding when and how to inform sample members that a survey requires involvement from the whole household. There is evidence for the positive benefits of transparency in communications,[13] but we must balance any potential benefits with the risk that informing people that all adults must take part increases the perceived burden of the request and could reduce motivation.

Two treatments were trialled: (1) the whole household nature of the survey was mentioned up front in the invitation letter, and (2) it is revealed to the first respondent once they have completed the household grid. This provides justification for the collection of contact details of these other adults. The 'any data' and 'complete data' address conversion rates for each treatment are shown in table 8.

Table 8: Response by 'all adults' introduction strategy

Treatment	Any data (95CI)	Complete data (95CI)
(i) Letter introduction	12.6% (11.6%-13.7%)	7.1% (6.4%-7.8%)
(ii) End-of-grid introduction to first respondent	12.3% (11.5%-13.2%)	6.4% (5.8%-7.1%)

There was virtually no treatment effect on the probability of collecting *any* data from the household but perhaps a hint of one in the higher 'complete data' rate among households where the 'all adults' design was revealed at the start. As with other treatments covered in this section, any effect is probably quite small, and this evidence at least is insufficient to reject the null hypothesis of no effect. Nevertheless, introducing the 'all adults' design in the letter is the more informative approach and does not appear to have depressed the response rate at all (possibly, the reverse). On that basis, it is the marginally preferable strategy.

As noted earlier, including this information in the letter does not appear to have led to any inflation in reported household size (to gain more incentives) so is 'safe' from that perspective. It is true that there is a potential confounding effect in that the treatment might have encouraged smaller households to respond at a higher rate, in which case some inflations of household size would have produced the null result we see. However, that is probably an over-convoluted explanation. It is more sensible to accept the simpler one that revealing the 'all adults' design up front does not lower the probability of collecting data from the household and nor does it reduce the accuracy of the household grid data.

[13] Related evidence on progress bars and telling respondents how long they have on a survey *e.g.* BETA (2018) 'Easy as pie: encouraging greater survey completion', Working Paper, Australian Department of Prime Minister and Cabinet, *https://behaviouraleconomics.pmc.gov.au/sites/default/files/projects/report-easy-as-pie.pdf*

Trial 7: Asking the first respondent to provide the email addresses of other adults in the household

In general, it is preferable to collect as many contact details as possible about each panel member in a longitudinal study.

Two treatments were trialled: (i) the first respondent in a household was asked to provide email addresses of other adults in the household so we could contact them that way as well as by written letter, and (ii) the first respondent was not asked for other people's email addresses. In fact, these email addresses were not used so this trial was really to test whether requesting this information from the first respondent would have an impact on the complete data rate. This first respondent will not necessarily agree to provide these email addresses (or even be able to do so) so the treatment effect is estimated as an intention-to-treat. The 'complete data' address conversion rates for each treatment are shown in table 9 (it could not influence the 'any data' rate so that is not shown).

Table 9: Response by email collection strategy

Treatment	Complete data (95CI)
(i) No 'other adult' email addresses requested of the first respondent	6.8% (6.1%-7.5%)
(ii) 'Other adult' email addresses requested of the first respondent	6.7% (6.2%-7.4%)

This treatment made virtually no difference to the 'complete data' rate. This suggests that it is 'safe' to request email addresses from the first respondent. The ability to contact the other adults by email as well as by letter should increase the W14 individual level module completion rate relative to the trial although we cannot say by how much.[14]

Within those multi-adult households where email addresses were requested, at least one was obtained in 68% of cases, and one was obtained for *all* other adults in 62% of cases. Overall, **an email address was obtained for 60% of enumerated other adults**. This email obtainment rate was lower the larger the household (49% among households with >=2 other adults, compared to 68% among households with only one other adult). Nevertheless, this procedure seems to have worked fairly well.

Trial 8: Including encouraging messages at intervals during the survey

The final treatment was designed to keep respondents on track to complete the questionnaire, and potentially to end it with a warmer feeling towards the study than would otherwise be the case (perhaps some might even advocate for the study among other adults

[14] We should acknowledge that it may not have a positive effect on every one of these 'other adults' (some may see contact by email after obtaining the address from a co-resident as a reason to reject the study).

in their household?). The idea was simply to insert encouraging messages ("e.g., "You're nearly there."), particularly after long or taxing sections of the survey, to encourage participants to keep going.

Two treatments were trialled: (1) respondents were shown a total of four different encouraging messages in different sections of the questionnaire, and (ii) respondents were not shown encouraging messages.

The first message was shown at the start of the household finance module: *"We're going to ask you some important questions about money. We know it's a sensitive subject so we want to remind you that all your answers will be kept safe and anonymous."* The second message was shown at the end of the household finance module *"Thank you for all the information you've shared so far. We will keep it safe."* The third was shown early in the self-completion module: *"We're going to ask some sensitive questions, so we just want to remind you that your responses will be anonymised, and no one will know it's you."* The final message was shown towards the end of the self-completion module: *"We appreciate some of those questions may have been difficult to answer, so thank you for sharing. Now keep going - you're not too far from the end of the survey."* The results of the trial are shown in table 10.

Table 10: Response by in-survey messaging strategy

Treatment	Complete data (95CI)
(i) No 'encouraging' messages	6.6% (6.0%-7.3%)
(ii) 'Encouraging' messages at intervals through the questionnaire	6.9% (6.2%-7.7%)

The messages appear to have made no significant difference to the complete data rate, although the observed rate was slightly higher than under the 'no messages' condition. To be fair, this trial probably did not have the power to detect a realistic-sized effect as the treatment can affect only those households that at least started the survey. In that context, a 0.3 percentage point improvement in the complete data rate would be appreciable. However, in this case, that gain is due to a random difference in the 'any data' rate between the two treatment arms. The conditional probability of complete data, given a household grid, was virtually identical: 54% with no encouraging messages, and 55% with encouraging messages.

Including encouraging messages is, more or less, a cost-free design decision but it is hard to separate out the concept from its execution. A different set of messages, or the same set presented differently, might lead to different results. Within the research team, there were some who found the messaging patronising, while others thought it reinforced the point that completing the survey is important and/or worthwhile. The null result from this trial might reflect a balance of positive and negative effects among responding households or there may have been no effects at all: it is impossible to say for sure. However, the net

impact is surely small, so the inclusion or exclusion of these messages is hardly critical for the W14 survey design.

Interactions between treatments

It is easy enough to think of some potential interactions between treatments, in which the presence of one particular treatment enhances or depresses the effect of another particular treatment when combined together. However, even a study of this scale only has the power to detect *very* large interaction effects. Other interaction effects of useful scale are likely to be obscured by random sampling error. To illustrate, for two-arm trials, we generally have the power (>=80% probability) to detect main effects on the 'any data' rate of 1.9%pts or greater, and on the 'complete data' rate of 1.3%pts. We only have equivalent power to detect interaction effects of at least *twice* this scale (i.e., 3.8%pts on the any data rate, and 2.6%pts on the complete data rate). A practical consequence is that even quite large observed interaction effects will have confidence intervals that include zero, and quite small ones will have an upper bound to its 95CI that is well above zero.

We did not design this study with firm hypotheses about any interaction effect, although some seem feasible: e.g., the effect of the incentive package might vary as a function of whether or not the 'all adults' design was cited in the letter, or whether or not a prenotification letter was sent. It also seems feasible that the willingness to supply email addresses for other adults in the household might vary depending on when the 'all adults' design had been introduced (in the letter or just after completing the grid).

The problem is that we would not expect very large effect sizes for any of these things. In practice, true interaction effects are unlikely to be much bigger than half the largest main effect in the model, which places a limit on their scale of around two thirds of the largest observed treatment effect evaluated in isolation. The largest observed treatment effect was for the prenotification letter - adding 3.2%pts to the 'any data' rate and 2.1%pts to the 'complete data' rate – suggesting that **interaction effects of one to two percentage points is the outer limit of what is reasonable to expect**.

Table 11 shows perhaps the strongest evidence for an interaction effect: that the early bird incentive model works a little better if a prenotification letter is sent than if it isn't (and especially if there are only two reminders). This makes sense as the prenotification letter works to prime (some) households so they are more ready to respond to the early bird offer when it comes. However, the confidence interval around the estimated effect of each treatment combination is so wide that zero effects cannot be discounted. For this reason, it is probably sensible to ignore the potential for interaction effects when estimating the effect of treatments in combination.

Table 11: Response by contact and incentive strategy (main effects + interaction model)

Treatment combination		Any data (95CI)	Complete data (95CI)	'Early bird' treatment effect, controlling for contact sequence (any data / complete data)
(i) £30 conditional incentive with 5-week deadline	(i) Invitation letter + 3 reminders	11.8% (10.2%-13.4%)	6.3% (5.2%-7.4%)	
	(ii) Prenotification letter + invitation letter + 3 reminders	14.3% (12.9%-15.7%)	8.5% (7.2%-9.7%)	
	(iii) Prenotification letter + invitation letter + 2 reminders	11.9% (10.5%-13.4%)	6.5% (5.4%-7.5%)	
(ii) 'Early bird': £20 conditional incentive with 5-week deadline plus extra £10 if respond within 2 weeks	(i) Invitation letter + 3 reminders	10.0% (8.7%-11.3%)	5.1% (4.3%-6.0%)	-1.8%pts / -1.2%pts
	(ii) Prenotification letter + invitation letter + 3 reminders	13.9% (12.2%-15.5%)	7.1% (6.0%-8.3%)	-0.4%pts / -1.4%pts
	(iii) Prenotification letter + invitation letter + 2 reminders	12.8% (11.3%-14.3%)	7.0% (5.9%-8.2%)	+0.9%pts / +0.5%pts

4. Conclusion: Recommended treatments for the W14 boost survey

Given the evidence above, the critical thing for the W14 boost survey is to ensure that:

(i) a prenotification letter is sent;
(ii) a third reminder is sent;
(iii) email addresses of 'other adults' are requested of the first respondent; and
(iv) (probably) an early-bird incentive bonus is offered that covers the web-only phase.

Beyond that, we should consider maximising (a) transparency and (b) logistical simplicity if this has no appreciable effect on the response rate. In this context, it would make sense to:

(v) include both the University of Essex and ISER logos (or at least one of them);
(vi) introduce the 'all adults' nature of the survey in the letter(s);
(vii) introduce the longitudinal nature of the survey in the letter(s);
(viii) do not include any post-it notes;
(ix) do not place encouraging messages in the survey script.

The web trial data cannot tell us what the 'any data' or 'complete data' rates would be under this design because (i) email addresses collected of other adults will actually be used in W14 instead of just collected, and (ii) the early bird incentive is likely to apply for the full web-only fieldwork phase rather than just the first two weeks. Neither of these were tested but both are likely to affect the response rates. *If we ignored these additional effects* - as well as the potential for any interaction effects – then **we might expect an 'any data' address conversion rate of 14.9% (95CI = 12.8%-17.0%) and a 'complete data' rate of 7.7% (95CI = 6.3%-9.1%) under this design**.[15]

[15] This makes the assumption that the higher response rates observed for the IP14 trial reflected its longitudinal character and not any other difference from the wider web trial. The variables in the 'main effects only' model cover design elements (i) to (iv) but not (v) to (ix) since the latter have been evaluated to have near-zero effects on the two response rate metrics. If the model is extended to cover design elements (v) to (ix), the 'any data' rate would increase slightly to 15.1% (95CI = 12.6%-17.6%) but the 'complete data' rate would decrease slightly to 7.3% (95CI = 5.7%-8.9%). These differences from the rates quoted in the main text almost certainly reflect random variation rather than any meaningful effects of extending the model's coverage.