



Understanding Society
Working Paper Series

No. 2019-03

May 2019

***Understanding Society* Innovation Panel Wave 11:
Results from Methodological Experiments**

Tarek Al Baghal (ed.)

**Contributors: Jonathan Burton³, Roxanne Connelly⁸, Mick P. Couper⁹,
Thomas F. Crossley¹⁰, Catherine De Vries¹, Vernon Gayle², Tim
Hanson⁴, Annette Jäckle³, Peter Lynn³, Nicole Martin⁵, Alice McGee⁴,
Christopher Playford¹¹, Stephen Pudney⁶, Maria Sobolewska⁵, Luke
Taylor⁴, Sandra Walzenbach³, Alexander Wenz⁷**

¹ University of Essex, ² University of Edinburgh, ³ Institute for Social and Economic Research,
University of Essex, ⁴ Kantar Public ⁵ University of Manchester, ⁶ University of Sheffield, ⁷ University
of Mannheim, ⁸ University of York, ⁹ University of Michigan, ¹⁰ European University Institute,
¹¹ University of Exeter

***Understanding Society* Innovation Panel Wave 11:
Results from Methodological Experiments**

Tarek Al Baghal (ed.)

Contributors: Jonathan Burton³, Roxanne Connelly⁸, Mick P. Couper⁹, Thomas F. Crossley¹⁰, Catherine De Vries¹, Vernon Gayle², Tim Hanson⁴, Annette Jäckle³, Peter Lynn³, Nicole Martin⁵, Alice McGee⁴, Christopher Playford¹¹, Stephen Pudney⁶, Maria Sobolewska⁵, Luke Taylor⁴, Sandra Walzenbach³, Alexander Wenz⁷

¹ University of Essex, ² University of Edinburgh, ³ Institute for Social and Economic Research, University of Essex, ⁴ Kantar Public ⁵ University of Manchester, ⁶ University of Sheffield, ⁷ University of Mannheim, ⁸ University of York, ⁹ University of Michigan, ¹⁰ European University Institute, ¹¹ University of Exeter

Non-technical summary

The *Understanding Society* survey includes what is known as an 'Innovation Panel' sample (IP). This sample of originally 1500 households is used to test different methods for conducting longitudinal surveys in order to produce the highest quality data. The results from the Innovation Panel provide evidence about the best way to conduct a longitudinal survey which is of relevance for all survey practitioners as well as influencing decisions made about how to conduct *Understanding Society*. This paper reports the experiments with the mixed-mode design and early results of the methodological tests carried out at wave 11 of the Innovation Panel in the spring and summer of 2018.

IP11 employed a mixed-mode design including an internet survey, and continued ongoing experiments on the impact of incentives. The mixed-mode experiment was extended at IP11, where half of the new refreshment sample added at this wave was initially invited to complete via the web. Previously, new samples were all invited to an in-person interview. As with prior waves, several other methodological experiments were also included in the survey. Several experiments were conducted on survey measurement, including using different versions of the frequently used EQ-5D scale; how 'Don't Know' responses were presented in the survey; youths' understanding of their parents' occupations; and people's attitudes towards immigration. Respondents were also asked to participate in a study to record their spending via a mobile app. Two additional experiments were included to explore linkage of external data sources to respondents' survey data. Requests were made to link respondents' electoral register data and for access to their HMRC records.

***Understanding Society* Innovation Panel Wave 11:
Results from Methodological Experiments**

Contributors: Jonathan Burton³, Roxanne Connelly⁸, Mick P. Couper⁹, Thomas F. Crossley¹⁰, Catherine De Vries¹, Vernon Gayle², Tim Hanson⁴, Annette Jäckle³, Peter Lynn³, Nicole Martin⁵, Alice McGee⁴, Christopher Playford¹¹, Stephen Pudney⁶, Maria Sobolewska⁵, Luke Taylor⁴, Sandra Walzenbach³, Alexander Wenz⁷

¹ University of Essex, ² University of Edinburgh, ³ Institute for Social and Economic Research, University of Essex, ⁴ Kantar Public ⁵ University of Manchester, ⁶ University of Sheffield, ⁷ University of Mannheim, ⁸ University of York, ⁹ University of Michigan, ¹⁰ European University Institute, ¹¹ University of Exeter

Abstract

This paper presents some preliminary findings from Wave 11 of the Innovation Panel (IP11) of *Understanding Society*: The UK Household Longitudinal Study. *Understanding Society* is a major panel survey in the UK. In May 2018, the eleventh wave of the Innovation Panel went into the field. IP11 used a mixed-mode design, using on-line interviews and face-to-face interviews. This paper describes the design of IP11, the experiments carried and the preliminary findings from early analysis of the data.

Key words: longitudinal, survey methodology, experimental design, respondent incentives, questionnaire design.

JEL classification: C80, C81, C83

Contact: Tarek Al Baghal (talbag@essex.ac.uk) Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK.

Table of Contents

1. Introduction	5
2. <i>Understanding Society</i> : the UKHLS	6
3. Innovation Panel Wave 11: Design.....	8
a. Call for experiments	9
b. Sample.....	10
c. Questionnaire design	11
d. Response rates	12
The Impact of Incentives.....	15
Longitudinal Response Outcomes	16
4. Experimentation in IP11	17
a. Experiments on the EQ-5D Scale (Stephen Pudney)	18
b. Do we know what to do with ‘Don’t Know’? (Tim Hanson, Alice McGee, Luke Taylor).....	21
c. Push-to-Web Experiments (Peter Lynn)	27
d. Mr. Chalk the Teacher: An experimental investigation of children’s consistency in reporting their parent’s occupations (Vernon Gayle, Roxanne Connelly, Christopher Playford).....	31
e. Consent to HMRC data linkage experiment (Annette Jäckle, Jonathan Burton, Mick P. Couper, Thomas F. Crossley, Sandra Walzenbach)	33
f. Does Competition over Public Services Decrease Support for Residency Rights of Immigrants? Evidence From the United Kingdom (Nicole Martin, Catherine De Vries)..	37
g. Spending Study 2: Mode of invitation experiment (Annette Jäckle, Alexander Wenz, Jonathan Burton, Mick P. Couper)	41

h. Evaluating opt-out consent for linkage to the electoral register (Nicole Martin, Maria Sobolewska)	43
References	45

1. Introduction

This paper presents early findings from the eleventh wave of the Innovation Panel (IP11) of *Understanding Society: The UK Household Longitudinal Study* (UKHLS). *Understanding Society* is a major panel survey for the UK. The first nine waves of data collection on the main sample have been completed, and tenth and eleventh waves are currently in the field. The data from the first eight waves of the main samples are available from the UK Data Archive, and the ninth will be available towards the end of 2018. Data from a nurse visit to collect bio-markers from the general population sample and the British Household Panel Survey (BHPS) are also available. Data for the first eleven waves of the Innovation Panel are available from the UK Data Service¹.

One of the features of *Understanding Society*, alongside the large sample size (40,000 households at Wave 1) and the ethnic minority boost sample and the collection of bio-markers, is the desire to be innovative. This has been a key element of the design of *Understanding Society* since it was first proposed. Part of this drive for innovation is embodied within the Innovation Panel (IP). This panel of almost 1500 households was first interviewed in the early months of 2008. The design in terms of the questionnaire content and sample following rules are modelled on *Understanding Society*. The IP is used for methodological testing and experimentation that would not be feasible on the main sample. The IP is used to test different fieldwork designs, new questions and new ways of asking existing questions.

The second wave of the Innovation Panel (IP2) was carried out in April-June 2009, the third wave (IP3) in April-June 2010 and the fourth wave in March-July 2011. The fourth wave of the Innovation Panel (IP4) included a refreshment sample of 465 responding households. In March 2012, IP5 was fielded, with part of the samples conducting the survey via the internet, while others continued in an interviewer-administered survey. Fieldwork for IP6 started in March 2013, repeating the design where some were first asked to complete the survey via the

¹ <http://discover.ukdataservice.ac.uk/series/?sn=2000053>

web option while others were approached by an interviewer only. The IP6 also included a mop-up follow-up phase with anyone not responding with contacts attempted by CATI or CAWI at the end of the fieldwork. IP7 started fieldwork in June 2015 and added 488 responding households as a refreshment sample. IP8 fieldwork started in summer 2015, and IP9 in May 2016. IP10 added another refreshment sample of 339 completing households, and was conducted beginning May 2017.

Working Papers which cover the experimentation carried out in all ten previous innovation panels are available from the *Understanding Society* website.² The data from the first ten waves of the Innovation Panel are held at the UK Data Service. This paper describes the design of IP11, the experiments carried and some preliminary findings from early analysis of the data. Section 2 outlines the main design features of *Understanding Society*. Section 3 describes the design and conduct of IP11. Section 4 then reports on the experiments carried at IP11.

2. *Understanding Society*: the UKHLS

Understanding Society is an initiative of the Economic and Social Research Council (ESRC) and is one of the major investments in social science in the UK. The study is managed by the Scientific Leadership Team (SLT), based at ISER at the University of Essex and includes topical experts (“Topic Champions”) from a number of institutions. The fieldwork and delivery of the survey data for the first five waves of the main samples were undertaken by NatCen Social Research (NatCen). Since Waves 6, Kantar Public has been the lead contractor. *Understanding Society* aims to be the largest survey of its kind in the world. The sample covers the whole of the UK, including Northern Ireland and the Highlands and Islands of Scotland. *Understanding Society* provides high quality, longitudinal survey data for academic and policy research across different disciplines. The use of geo-coded linked data enables greater research on neighbourhood and area effects, whilst the introduction of

² <https://www.understandingsociety.ac.uk/research/publications/working-papers>

bio-markers and physical measurements (Waves 2 and 3) opens up the survey to health analysts.

The design of the main-stage of *Understanding Society* is similar to that of the British Household Panel Survey (BHPS) and other national panels around the world. In the first wave of data collection, a sample of addresses was issued. Up to three dwelling units at each address were randomly selected, and then up to three households within each dwelling unit were randomly selected. Sample households were then contacted by NatCen interviewers and the membership of the household enumerated. Those aged 16 or over were eligible for a full adult interview, whilst those aged 10-15 were eligible for a youth self-completion. The adult interviews were conducted using computer-assisted personal interviewing (CAPI) using laptops running the questionnaire in Blaise software. Adults who participated in *Understanding Society* were also asked to complete a self-completion questionnaire, in which questions thought to be more sensitive were placed. The adult self-completions at Waves 1 and 2, and the youth self-completions, were paper questionnaires. From Wave 3 onwards the adult self-completion instrument was integrated into the interviewing instrument and the respondent used the interviewer's lap-top to complete that portion of the questionnaire themselves (Computer-Assisted Self-Interviewing, CASI). For the first seven waves, surveys of continuing sample members were interviewer-administered. Before Wave 7 was issued, a random 20% of households were designated as ring-fenced face-to-face and would only be issued to CAPI first. Except for this ring-fenced sample, households that had not responded at Wave 6 were issued at Wave 7 to a sequential mixed mode design in which adults in those households were initially invited to participate online, and then those who did not take part online were issued to interviewers ("WEB"). At Wave 8, the proportion of households that were issued web-first increased to 40%; this included households that had not participated at Wave 7, as well as those households who were predicted to be most likely to complete online – except for those who were in the ring-fenced sample. The remaining 60% of households were issued to interviewers, with non-respondents at the reissue stage invited to complete online ("F2F"). At Wave 9, the proportion of households issued web-first increased to 60%, and at Waves 10 and 11 increased again to 70%.

In between each wave of data collection, sample members are sent short reports of early findings from the survey, and a change-of-address card, to allow them to inform ISER of any change in their address and contact details. Before each sample month is issued to field for a new wave, each adult is sent a letter which informs them about the new wave of a survey, includes a token of appreciation in the form of a gift voucher and also includes a change-of-address card. Interviewers then attempt to contact households and enumerate them, getting information of any new entrants into the household and the location of anyone who has moved from the household. New entrants are eligible for inclusion in the household. Those who move, within the UK, are traced and interviewed at their new address. Those people living with the sample member are also temporarily eligible for interview. More information about the sampling design of *Understanding Society* is available in Lynn (2009).³ From Wave 2, the BHPS sample has been incorporated into the *Understanding Society* sample. The BHPS sample is interviewed in the first half of each wave.

3. Innovation Panel Wave 11: Design

IP11 was comprised of five samples: the original sample from IP1, and refreshment samples taken at IP4, IP7 IP10, and IP11. IP11 employed a mixed-mode design, which started in IP5 has been used in each subsequent wave. Starting at IP5, the modes which were mixed were on-line (CAWI) and face-to-face (CAPI) interviewing. In IP5, a random selection of two-thirds of households was allocated to the mixed-mode design (“WEB”) with the remaining third of households allocated directly to face-to-face interviewers (“F2F”). This sample allocation has been maintained at each wave. However, starting at IP8 subgroup of households with a very low propensity to respond via the web in in the CAWI condition was assigned to CAPI to begin fieldwork. Very low web propensity was determined by modelling web-completion using previous IP data. Unlike previous refreshment samples, the IP11 refreshment sampled households were allocated to either face-to-face (F2F) or web-first (WEB) recruitment, described in experimental results below, in Section 4.c. In England and Wales, face-to-face

³<https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2009-01.pdf>

interviewing assignments were evenly split between Kantar (the lead contractor) and NatCen. Kantar conducted all the face-to-face interviewing assignments in Scotland.

Initially, advance letters were sent to adults in the WEB group which included a URL and a unique log-in code. The IP11 refreshment sample households did not receive individual personalized letters; rather these received a household-wide invitation via a letter in the mail. Adults in the WEB group for whom we had an email address (none in the IP11 refreshment sample) were also sent an email which included a link which could be clicked through to the web-site. There were two email reminders for adults with an email address who had not yet completed their interview on-line. A reminder letter was then sent to all households in the WEB group who had not completed their interview via the web one week after the initial invitation. For the IP11 refreshment sample allocated to the WEB condition, a second reminder letter was sent to non-responding households two weeks after the initial invitation. A soft launch, comprising 10% of all WEB-first households, took place on 24 May, with the full launch for the remaining WEB-first households on 30 May.

After four weeks of the CAWI survey being in the field, CAPI fieldwork started for IP11. CAPI interviews were attempted for all adults in the CAPI-only design and those who had not completed their WEB interview in the mixed-mode design. Those not responding to the WEB survey during the initial CAWI field period were allocated to face-to-face interviewers, but could still enter the web survey instead if they desired. Adults who had started their interview on-line, but not reached the 'partial interview' marker, were issued to face-to-face interviewers. The interviewers were able to re-start the interview at the place at which the respondent had stopped. Total fieldwork lasted 20 weeks. Overall, the WEB-only period took place in 24th May to June 26th June, while face-to-face fieldwork started June 27th and continued until 9th October. The mop-up follow-up phase attempted interviews with those not responding in both the WEB and F2F versions, through CAPI, CATI or CAWI versions of the survey. This final phase ran from 10th October until 21st October.

a Call for experiments

IP11 was the ninth time the Innovation Panel was open for researchers outside the scientific team of *Understanding Society* to propose experiments. A public call for proposals was made 6th February with a deadline of 30th March. Twenty-nine proposals were received with seven

being accepted. In addition to these seven experiments, an additional experiment was continued from IP10 (on electoral consent linkage, Section 4.g), for a total of eight experiments carried in IP11. The initial proposals were reviewed by a panel which included two ISER-based members of the *Understanding Society* scientific leadership team, and two members of the Methodology Advisory Committee to *Understanding Society* who were external to ISER. In addition to those experiments which were accepted through the public call, there were two continuing core experiments which the Understanding Society senior leadership team wanted to run. These core experiments are the mixed-mode design and the main incentives experiment.

b. Sample

There were five samples issued at IP11: the original sample from IP1 and refreshment samples issued at IP4, IP7, IP10, and IP11. Samples other than the IP11 refreshment sample were comprised of those households who had responded at IP10, plus some households which had not responded at IP10. Households which had adamantly refused or were deemed to be mentally or physically incapable of giving an interview were withdrawn from the sample. There were 732 original sample households, 345 IP4 refreshment sample households, 428 IP7 refreshment sample households, 350 IP10 refreshment sample households, and 2532 IP11 refreshment sample households issued. There were 4387 total sample households issued at IP11. All of the households were originally selected from the Postcode Address File (PAF) using the same methods.⁴

As noted above around two-thirds of the original, IP4, IP7, and IP10 refreshment samples were allocated to the mixed-mode design which was initiated at IP5, and which has been maintained all subsequent waves, including IP11. Sample members would be approached by letter and email (where possible) to complete their interview on-line. The IP11 refreshment sample was also allocated to a mixed-mode design, but at a 2/3 F2F and 1/3 WEB allocation. This contrasts with prior refreshment samples, which were all allocated to F2F in its first

⁴ See Lynn, P. (2009). Sample Design for Understanding Society *Understanding Society Working Paper Series No. 2009 – 01* at <https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2009-01>

year of data collection. As noted, some households that were deemed to have a very low propensity to respond via WEB were assigned directly to CAPI to begin fieldwork (in samples other than IP11 refreshment sample, where no prior information existed). The table below shows the allocation to mode design by sample type for those included in the issued original, IP4, and IP7 refreshment samples in IP10.

Table 1: Allocation to mode design by sample type

	Original	IP4 Refreshment	IP7 Refreshment	IP10 Refreshment	IP11 Refreshment
CAPI only	259 35.4%	133 38.6%	140 32.7%	117 33.4%	1690 66.8%
Mixed-mode (CAWI+CAPI)	473 64.6%	212 61.5%	288 67.3%	233 66.6%	842 33.3%
Total	732	345	428	350	2532

c. Questionnaire design

The questionnaire at IP11 followed the standard format used in the previous Innovation Panels as well as the main-stage of *Understanding Society*. The questionnaires used at IP11 are available from the *Understanding Society* website.⁵ The interview included the following sections with the corresponding target times for each:

- Household roster and household questionnaire: 15 minutes per household
- Individual questionnaire: average 31 minutes for each person aged 16 or over
- Adult self-completion: around 9 minutes, computer self-administered interview (CASI)
- Youth self-completion: 10 minutes for each child aged 10-15 years
- Proxy questionnaire: 10 minutes for adults ages 16 or over who are not able to be interviewed.

⁵ <https://www.understandingsociety.ac.uk/documentation/innovation-panel/questionnaires>

There were some changes made to the questionnaire to enable participants to complete it online at IP5 when the web design was first introduced, and can be described more in-depth in the working paper containing results from the experiments in IP5.⁶ Briefly, the changes made to the questionnaire are as follows. Questions were reworded as needed to include interviewer instructions that may clarify the definition of the question. Text was altered to be more participant-focused rather than interviewer-focused. The first person in the household to log in to the web survey would be asked to complete the household enumeration. A question about who was responsible for paying household bills was included; the person or people indicated as responsible were routed first to the household questionnaire and then to the individual questionnaire.

If a participant had started to answer their questionnaire and left the computer for 10 minutes, they were automatically logged out. The participant was able to log back in using the same process as they had originally logged in, and they would be taken to the place that they had left the interview. This also applies to those who had closed down the browser mid-interview. A 'partial interview' marker was put into place about two-thirds of the way through the interview, after the benefits section. If a participant reached this stage, the interview was considered to be a 'partial interview'. They could log back in and complete if they wanted, but otherwise they were not contacted by an interviewer. If the participant had not reached this marker before closing down the browser, they were sent an email overnight which thanked them for their work so far and encouraged them to complete the survey, giving them the URL to click through to the survey. Again, they would start at the point where they had left off. In addition, those who had started but not reached the partial interview marker were, after the initial two weeks, issued to face-to-face interviewers who would be able to finish the survey with them, from where they had left off.

d. Response rates

This section sets out the response rates for IP11 as a whole. The issued sample at the eleventh wave consisted of 1579 households that had responded to IP10, 2532 IP11 refreshment sample households, and 279 households that had not responded at IP10, but had

⁶<https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2013-06>

at some prior wave(s). Fieldwork for all sample members was split by CAPI-only and mixed-mode allocations. Table 2 displays the household-level response at IP11 for the continuing samples by CAPI-only and mixed-mode conditions and the overall total response. The lower panel displays individual response rate for each. For each cell, the percent is reported above the number of units the percent represents, in italics. The total number of eligible sampled units is in the Total rows, in bold.

Table 2. Household and Individual Response Outcomes for Continuing Samples by Mode Design, IP11

	Original		IP4 Refreshment		IP7 Refreshment		IP10 Refreshment		Total
<i>Household RR</i>	F2F	MM	F2F	MM	F2F	MM	F2F	MM	
Complete HH	58.9% <i>145</i>	58.3% <i>268</i>	50.0% <i>66</i>	57.7% <i>120</i>	47.8% <i>64</i>	52.7% <i>147</i>	35.7% <i>41</i>	44.6% <i>103</i>	52.9% <i>954</i>
Partial HH	21.5% <i>53</i>	19.4% <i>89</i>	17.4% <i>23</i>	18.3% <i>38</i>	23.1% <i>31</i>	19.7% <i>55</i>	26.1% <i>30</i>	21.2% <i>49</i>	20.4% <i>368</i>
Total Responding HH	80.5% <i>198</i>	77.6% <i>357</i>	67.4% <i>89</i>	76.0% <i>158</i>	70.9% <i>95</i>	27.6% <i>77</i>	61.7% <i>71</i>	65.8% <i>152</i>	73.2% <i>1805</i>
Nonresponding HH	19.5% <i>48</i>	22.4% <i>103</i>	32.3% <i>43</i>	24.0% <i>50</i>	29.1% <i>39</i>	27.6% <i>77</i>	38.3% <i>44</i>	34.2% <i>79</i>	26.8% <i>483</i>
Total HH	246	460	132	208	134	279	115	231	1805
<i>Conditional Individual RR</i>	F2F	MM	F2F	MM	F2F	MM	F2F	MM	
Responding individuals	80.8% <i>395</i>	83.2% <i>628</i>	81.0% <i>141</i>	84.2% <i>271</i>	78.5% <i>150</i>	80.8% <i>329</i>	64.5% <i>91</i>	77.4% <i>229</i>	80.5% <i>2134</i>
Nonresponding individuals	19.2% <i>70</i>	16.8% <i>138</i>	19.0% <i>33</i>	15.8% <i>51</i>	21.5% <i>41</i>	19.2% <i>78</i>	35.5% <i>50</i>	22.6% <i>67</i>	19.5% <i>517</i>
Total Ind.	365	755	174	322	191	407	141	295	2651

There were 1805 interviewed households from the continuing samples, for a 73.2% overall household response rate. Within these households, 2134 people were interviewed, for a conditional individual response rate of 80.5%.

Table 3 shows the household-level and individual-level response at IP11 for the IP11 refreshment sample. The eleventh wave was the initial wave for this sample, and the percentage allocated to each mode differed from other samples. For the IP11 refreshment

sample, 575 households were surveyed, a 24.4% response rate. Of all of the enumerated individuals in these households, 762 were interviewed, equalling a 73.6% conditional individual response rate.

Table 3. Household and Individual Response Outcomes for IP11 Refreshment sample, IP11

<i>Household RR</i>	F2F	MM	Total
Complete HH	16.9%	12.2%	15.3%
	266	96	362
Partial HH	8.4%	10.2%	9.0%
	133	80	213
<hr/>			
Total Responding HH	25.3%	22.4%	24.4%
	399	176	575
Nonresponding HH	74.7%	77.6%	75.6%
	1176	609	1785
Total HH	1575	785	2365
<hr/>			
<i>Conditional Individual RR</i>			
Responding individuals	76.5%	66.8%	73.6%
	555	207	762
Nonresponding individuals	23.6%	33.2%	26.6%
	171	103	274
Total Individuals	640	310	1036

Given the mixed-mode design used, not all individuals responded in the same mode. Further, at IP11 the mop-up period was again used, where non-responding units in all the samples were contacted and could respond via the web or telephone regardless of the allocated mode design. Only one person responded via the telephone at IP11. Table 4 shows the mode of completion for individuals by mode condition and total overall at IP11 including the mop-up phase. The continuing samples (Original, IP4 refreshment, IP7 refreshment, and IP10 refreshment) are combined, with the IP11 refreshment sample presented separately given the initial invitation and difference in design for this group this wave.

Table 4. Mode of Response, IP11

<i>Responding Mode</i>	Continuing Samples		IP11 Refreshment		Total
	F2F	MM	F2F	MM	
Face-to-Face	91.6% 620	20.2% 294	98.2% 545	66.2% 137	55.1% 1596
Telephone	-- 0	0.1% 1	-- 0	-- 0	0.03% 1
Web	8.4% 57	79.8% 1162	1.8% 10	33.8% 70	44.9% 1299
Total Ind.	677	1457	555	1336	2896

Starting in IP8, it was possible to access the web survey using any internet-enabled device. In previous waves, smartphones were blocked from accessing the survey, although tablets could access the questionnaire. A number of variables were captured about the device the survey was accessed with, including what type of device was used, the operating system, the device model, the browser used, browser version, and screen resolution. These variables are now available in the IP from the seventh wave as `w_deviceused` `w_deviceos` `w_devicemodel` `w_browserused` `w_browserversion` `w_screenresolution` in the file `w_indresp_ip`. The distribution of devices among web respondents used across all samples in IP11 is presented in Table 5.

Table 5. Device Used, Web Respondents, IP11

IP11 Web Respondents	
PC/Laptop	56.7% 736
Large Tablet	17.9% 233
Small/Medium Tablet	11.7% 152
Smartphone	13.7% 178
Total Web Respondents	1299

The Impact of Incentives

Most continuing sample members received the same incentive at IP11 as they had done at IP10. In consequence, there were again three experimental groups amongst the continuing

mixed mode samples (£10 unconditional, with or without an additional £20 conditional on participation online, £30 unconditional), three experimental groups amongst the IP7 refreshment sample in the F2F condition (£10, £20 or £30, unconditional). All IP11 sample members received £10 unconditionally, with individuals in the mixed-mode design receiving an extra £15 conditional on their completion of the web survey within the first three weeks. The upper panel of Table 6 presents total household response rates (including complete and partial response) by incentive type, excluding the IP11 refreshment sample, which had a somewhat different incentive structure, and was responding for the first time. The lower panel of Table 6 presents the impact of incentives on the IP11 refreshment sample.

Table 6. Household Response Rate by Incentive Type, IP11

<i>Continuing Samples</i>	HH Response
£10 Unconditional	70.2% 550
£20 Unconditional	65.2% 90
£10 Unconditional +£20 for individual	78.5% 233
£30 Unconditional	81.8% 306
<hr/>	
<i>IP11 Refreshment</i>	
£10 Unconditional	23.8% 399
£10 Unconditional +£15 for web in period	30.6% 176

Longitudinal Response Outcomes

The individual re-interview rate is an important outcome in a longitudinal survey, since analyses require pairs of observations to measure change. Re-interview rates are calculated as the percentage of eligible units responding at later waves who were also surveyed at the initial wave. For those in the original sample, the percentage is predicated on response at IP1, while the fourth wave is the initial wave for the IP4 refreshment sample, the seventh wave was the first for IP7, and tenth wave being the first for the IP10 refreshment sample.

Table 7 presents the longitudinal individual re-interview rates for the original sample (for IP2-IP10), the IP4 refreshment sample (for IP5-IP10), IP7 (for IP8-IP10), and IP10 (for IP11). For each cell, the percent is reported above the number of individuals the percent represents, in italics.

Table 7. Longitudinal re-interview rates

	IP2	IP3	IP4	IP5	IP6	IP7	IP8	IP9	IP10	IP11
Original Sample	69.3% <i>1654</i>	60.6% <i>1442</i>	54.7% <i>1270</i>	45.9% <i>1095</i>	45.9% <i>1100</i>	38.4% <i>917</i>	36.2% <i>867</i>	35.8% <i>814</i>	31.2% <i>746</i>	28.9% <i>691</i>
IP4 Refreshment	-	-	-	82.0% <i>586</i>	76.8% <i>554</i>	62.1% <i>447</i>	58.8% <i>423</i>	58.7% <i>396</i>	48.4% <i>350</i>	44.5% <i>321</i>
IP7 Refreshment							79.2% <i>520</i>	82.7% <i>487</i>	61.8% <i>404</i>	56.9% <i>371</i>
IP10 Refreshment										82.1% <i>422</i>

As with any longitudinal study, there has been attrition at each wave, decreasing the overall numbers for each sample. At IP11, 691 individuals from the original sample who responded at IP1 were successfully interviewed, representing a 28.9% re-interview rate. For the IP4 refreshment sample, the IP11 was their seventh wave and 321 responded, for a 44.5% re-interview rate. IP11 was the fourth wave for the IP7 refreshment sample, with 371 responses for a 56.9% re-interview rate. The IP10 refreshment sample was asked to participate for only the second time at IP11, and had 422 completes, an 82.1% re-interview rate.

4. Experimentation in IP11

There were a number of experiments carried on IP11, with one covering fieldwork procedures and the remainder examining measurement in the questionnaire. This section outlines the experiments carried at IP11; briefly explaining the reasons for carrying them, describing the design of the experiment and giving an indication as to the initial results from early analysis of the data. The analyses in this working paper were based on a preliminary data-set which contained all cases but did not have weights or derived variables. The authors and proposers of the experiment of each sub-section below are given in the heading.

a. Experiments on the EQ-5D Scale (Stephen Pudney)

EQ-5D is a 5-item questionnaire module designed to measure health-related quality of life. It is extremely important for policy purposes, since it is used to measure health benefits as part of many cost-effectiveness studies that provide evidence for decisions on medical technologies which are made by the National Institute for Health and Care Excellence (NICE) in England and similar policy bodies in other countries. The original version of EQ-5D featured five items relating to: mobility; self-care; ability to perform usual activities; pain; and anxiety or depression. Each item had a response scale with three levels, indicating no difficulties, some difficulties, or extreme difficulties.

EQ-5D has been redesigned (Herdman et al 2011) in a more detailed form (EQ-5D-5L) to increase its sensitivity, with 5-level response scales indicating: no / slight / moderate / severe / extreme problems. Much existing cost-effectiveness evidence is based on the older version (EQ-5D-3L) but, increasingly, trials of medical interventions are using the 5L version. Statistical prediction of 5L outcomes from 3L or *vice versa* are widely used to translate results from one basis to another (a procedure known as statistical mapping). Statistical mapping requires the availability of a multi-instrument (MI) survey that contains both the EQ-5D-3L and EQ-5D-5L instruments for the same group of respondents (see Hernández and Pudney (2017) and Wailoo et al. (2018) for recent examples of mapping between 3L and 5L)..

Our randomised experiment in the UKHLS Innovation Panel is designed to investigate two primary issues: (1) Does the repetition of the five EQ-5D questions in very similar 3L and 5L forms cause any distortion in the data? (By distortion, we mean that the distribution of responses is systematically different from the distribution of responses that would have been obtained in a survey that asked EQ-5D only once.) (2) Does ordering of the 3L and 5L instruments affect the distribution of responses?

The experimental design partitions the sample of households into three equal-sized randomly-selected groups:

- Group A receive EQ-5D-5L approximately half way through the interview and EQ-5D-3L three-quarters of the way through.

- Group B receive EQ-5D-3L approximately half way through the interview and EQ-5D-5L three-quarters of the way through.
- Group C receive only EQ-5D-5L, three-quarters of the way through.

In addition, interviewers were asked to make an assessment of the reactions of respondents in groups A and B to receiving repeated EQ-5D instruments in face-to-face mode, distinguishing between those who appeared not to notice the repetition, those who noticed but appeared unconcerned, and those who expressed dissatisfaction or other concern about the repetition. Among the 861 respondents who received both 3L and 5L by CASI and for whom there was complete personal data, those reactions were independent of gender and the 3L/5L ordering, but there was a large age difference in the reactions. Older people (over 60) were more than twice as likely as younger people (under-35) to express awareness (11% vs 24%) and concern over the repetition (12% vs 27%).

Table 8 reports the results of χ^2 tests of the hypothesis of equality of response distributions between pairs of treatment groups, separately for each of the five EQ-5D domains. In four of the five health domains, there are strongly significant differences in the response distributions between groups A and B, both in terms of their 3L and 5L responses. Consequently, the ordering of the two versions of EQ-5D does significantly influence the distribution of responses. Significant differences are less evident in the comparisons with group C, but groups A and B both display significant differences from group C in the pain dimension.

Table 8 χ^2 test statistics for equality of response distributions across treatment groups by domain of EQ-5D-5L					
Comparison	Health domain				
	Mobility	Self-care	Activities	Pain	Anxiety
<i>Equality of 3L response distributions: $\chi^2(2)$ statistic</i>					
group A vs group B	6.36**	9.90***	6.87**	8.64**	3.81
<i>Equality of 5L response distributions: $\chi^2(4)$ statistic</i>					
group A vs group B	11.51**	10.11**	5.02	47.31***	8.88*
group A vs group C	11.83**	5.58	7.55	11.74**	4.54
group B vs group C	5.97	7.67	2.94	19.40***	10.22**

† Sample of first-interviewed member in each household: all = all group members; LSI = group members reporting a long-standing illness or disability. Statistical significance: * = 10%, ** = 5%, *** = 1%.

As a simple aid to interpreting these significant differences, Table 9 shows the results of a regression analysis of the “misery index” formed by summing the numerical responses, giving an index ranging from 5 to 15 for 3L and 5 to 25 for 5L. The regression models include simple demographic characteristics as well as experimental group identifiers. The results show that participants in group B (5L preceded by 3L) have on average significantly lower “misery” (*i.e.* better reported health) and that the difference is comparable in size with the gender difference and age gradient in health.

These large significant effects suggest a need for caution in using results from multi-instrument surveys to estimate statistical mapping models – there may be significant distortions from the impact of question repetition on data quality. A paper by Hernández *et al.* (2019) gives more detail on these findings and considers the implications for policy analysis.

Table 9 Regression models for the “misery” index		
Covariate	3L misery index	5L misery index
Group A		0.143
Group B	-0.236***	-0.485***
Age 35-59	0.352***	0.335***
Age over 60	0.531***	0.999***
Female	0.202***	0.312***
Statistical significance: * = 10%, ** = 5%, *** = 1%. Standard errors adjusted for clustering by household.		

b. Do we know what to do with ‘Don’t Know’? (Tim Hanson, Alice McGee, Luke Taylor)

This experiment compares three treatments of ‘Don’t know’ (DK) response codes within the self-completion (CAWI and CASI) questionnaire.

Background

A wealth of evidence exists on the topic of DK response options including arguments around whether they should be shown as explicit options within questionnaires. Evidence is mixed and as a result, treatment of DK codes has been the subject of extensive debate among researchers.

With surveys increasingly moving to mixed-mode designs, comparability between modes is a growing concern. In an interviewer-administered context the interviewer can code a DK response where this is offered spontaneously. This approach cannot be replicated online although, in an attempt to move as close as possible to this design, some online surveys adopt a ‘hidden DK’ approach, where DK codes appear only where the respondent attempts to move on without selecting a response. This approach is currently taken on Understanding Society. However, usability testing in 2016/17 uncovered issues faced by respondents. This included cases where respondents were unaware of how to code a DK response and so selected an alternative option simply to allow them to move on in the survey. Concerns over these issues contributed towards the rationale for this experiment.

Experiment design

This experiment seeks to address four key research questions:

1. Does varying the treatment of DK codes produce different levels of DK response?
2. Does this vary between different types of question?
3. Is there any impact on the distribution of ‘substantive’ responses?
4. Does making DK less visible result in more ‘non-attitudes’⁷?

Three different treatments of DK responses were included:

1. ‘Hidden DK’: where DK appears if respondent attempts to move on without selecting a response
2. ‘Prompted DK’: As 1, but with a prompt on each screen for what to do if don’t know or don’t want to answer
3. ‘Offered DK’: DK code included as part of main list (always visible)

Respondents were randomly allocated to one of the three groups. A total of 2,605 respondents were included in the experiment, distributed broadly evenly across the three groups.

The experiment was conducted on 26 self-completion questions within the CAPI and CAWI questionnaires; for the CAPI interviews these fell in the self-completion part of the interview. Twenty-four of these questions collected self-assessed health measures and the remaining two were attitudinal questions on topics where salience and knowledge were expected to be low, specifically attitudes towards nuclear energy and trust in the United Nations. We also asked respondents to self-rate their knowledge about these two areas to allow us to explore whether hiding DK options leads to the collection of non-attitudes. In order to capture this, we included an open follow-up ‘clarification’ question where there was possible contradiction between self-reported levels of knowledge and whether an attitude was given.

⁷ ‘Non-attitudes’ refers to cases where a respondent’s ‘true’ answer is DK but, as they cannot easily code this, they select from one of the other (i.e. non-DK) response codes instead.

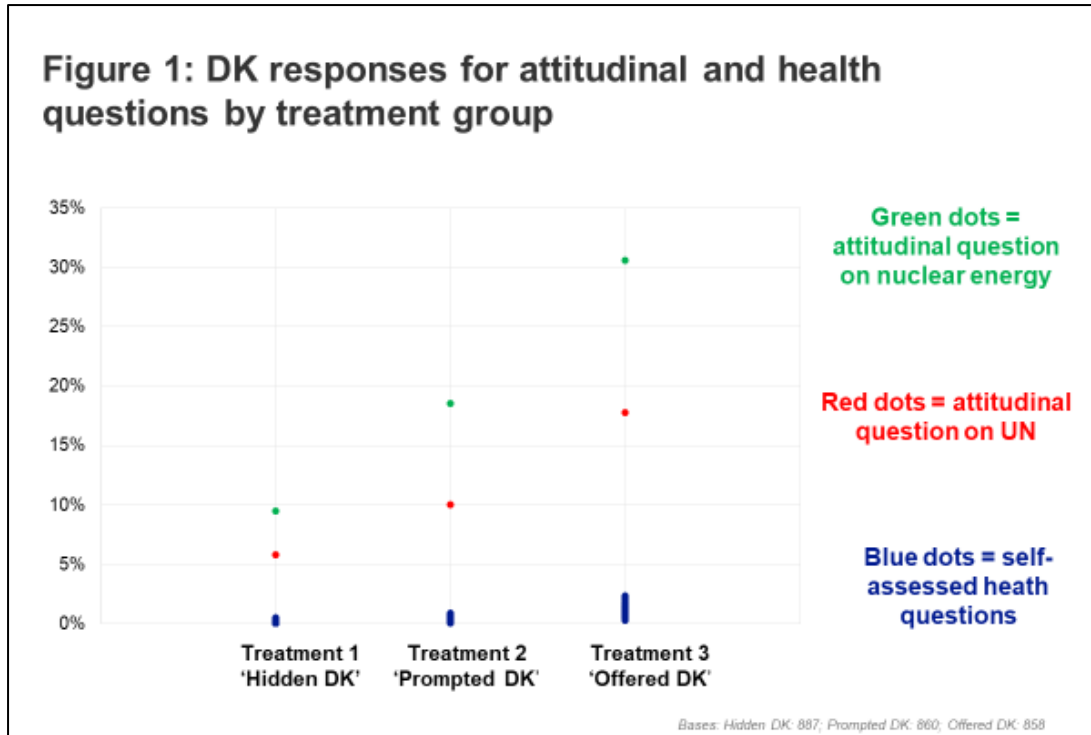
Results

Variation of DK response levels between treatment groups and different types of question

Analysis showed that treatment 3 ('Offered DK') consistently elicited a higher proportion of DK responses than treatment 1 ('Hidden DK'). There was a higher DK rate at 16 of the 24 self-assessed health measures and at both attitudinal questions ($p \leq 0.01$). There was also a significant difference when comparing treatment 2 ('Prompted DK') with treatment 1 for both attitudinal questions.

The differences at self-assessed health measures tended to be small (typically 1-2 percentage points) but overall levels of DK were very low (below 3%) for all three formats at these questions.

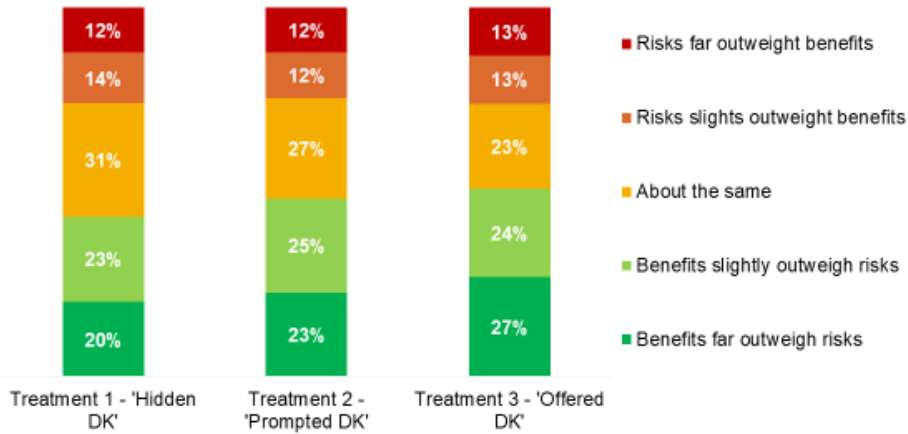
As Figure 1 shows, the differences were particularly pronounced for the two attitudinal questions on low-salience issues. When asked about the benefits and risks of nuclear energy, only 9% of those exposed to treatment 1 ('Hidden DK') answered DK, compared with 18% for treatment 2 ('Prompted DK') and 31% for treatment 3 ('Offered DK'). The pattern for the question about trust in the United Nations was very similar (Treatment 1, 6%; Treatment 2, 10%; Treatment 3, 18%).



Impact of treatment of DK codes on 'substantive' responses

With regard to the impact treatment of DK codes have on 'substantive' responses, our results show that, where DK codes are hidden (treatments 1 and 2), the middle category ('Benefits and risks are about the same') or the mid-point of the scale (point 5 on a 0-10 point scale) were selected in a greater number of cases. Figure 2 shows that when asked about the benefits and risks of nuclear energy the middle category was selected by 31% of respondents in treatment 1, 27% in treatment 2 and 23% in treatment 3. Figure 2 also shows that the proportion thinking the benefits of nuclear energy outweigh the risks differs between the three treatments (once DK responses are removed).

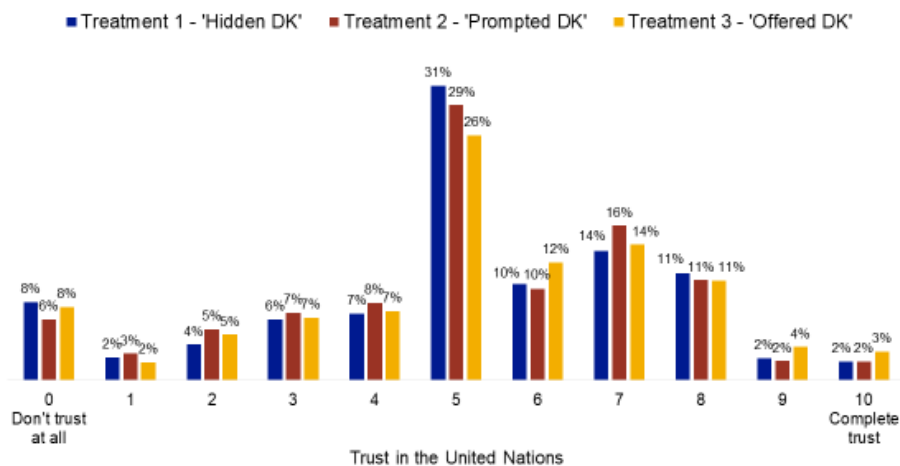
Figure 2: Attitudes towards Nuclear Energy by treatment group



Bases: Hidden DK: 785; Prompted DK: 666; Offered DK: 561

Figure 3 shows that, as with attitudes towards nuclear energy, very similar proportions of respondents selected the mid-point (point 5 on a scale of 0-10) when asked about their level of trust in the United Nations (treatment 1, 31%; treatment 2, 29% and treatment 3, 26%).

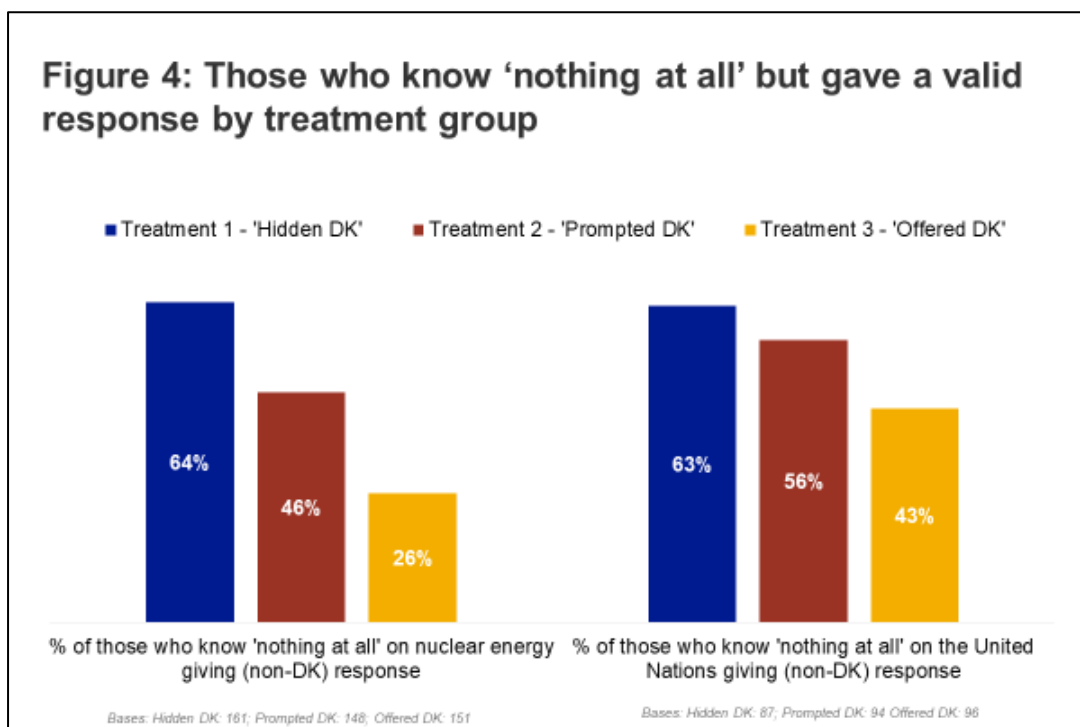
Figure 3: Attitudes towards United Nations by treatment group



Bases: Hidden DK: 824; Prompted DK: 745; Offered DK: 683

Impact of treatment of DK codes on levels of ‘non-attitudes’

Under treatment 1, 64% of those that reported knowing ‘nothing at all’ about nuclear energy provided a valid (i.e. non-DK) answer at the benefits and risks question; this fell to 46% for treatment 2, and 26% for treatment 3. The corresponding figures for the United Nations questions were: 63% (treatment 1), 56% (treatment 2) and 43% (treatment 3) respectively, showing a similar, although less pronounced, pattern. This suggests that ‘hiding’ a DK code may encourage the reporting of non-attitudes for these questions.



These findings build on the argument put forward by Converse (1976) and supported by findings from our usability testing, that respondents report ‘non-attitudes’ where DK options are not offered (Converse 1976).

Conclusions

In conclusion, there is clear evidence that varying the treatment of DK codes can impact heavily on DK rates and that this difference is most pronounced for low-salience attitudinal

questions. Whether DK codes are offered, prompted, or hidden can also impact on overall response distributions. While the implications of these results need further consideration, they should encourage researchers to more carefully consider the impact of different treatment of DK codes for future studies.

c. Push-to-Web Experiments (Peter Lynn)

Part of the IP11 refreshment sample was initially invited to participate online. This was something of a departure for Understanding Society, as all previous new samples on both the main survey and the IP had always been approached initially face-to-face. If successful, pushing (some) participants online immediately would be likely to save substantial survey costs, compared to a universal face-to-face approach. However, the likely success of the approach was unknown. Consequently it was decided to test the approach experimentally, with a random subsample approached web-first and the remainder approached CAPI-first, as had been done for previous refreshment samples.

Furthermore, it is not yet clear how best to introduce the invitation to complete a survey online in the complex context of a mixed-mode household survey with an address-based sample: the development of push-to-web methodology is still in its infancy (Dillman 2017). In particular, when the available sampling frame is just a list of addresses, the number of eligible people at each address, let alone their names, is unknown at the time of sending out the invitations. It is therefore not possible to send a personal invite to each household member. Instead, a single letter is sent to the household. There are a number of possible advantages and disadvantages of explaining upfront that all adult household members will be requested to participate. Similarly, there are potential advantages and disadvantages to notifying sample households from the outset that if they do not participate online, an interviewer will call at their address to request a CAPI interview. To test alternative ways of explaining these issues to sample households, an experiment was carried out within the web-first part of the IP11 refreshment sample, with two crossed factors: a) whether sample households were told in the invitation letter that there would be a CAPI follow-up if they did not participate online; and b) whether they were told in the invitation letter that all adult household members would be invited to participate.

In summary, the experiments involved the following treatments. Note that experiments #2 and #3 were orthogonal, involving the same set of 842 households and therefore around 210 in each of the four combinations of the two sets of experimental treatments. Experiments #2 and #3 were carried out on the households constituting group 2 of experiment #1.

Push-to-web experiment #1: Initial mode of data collection

Group 1 (n = 1,690 households), **CAPI-first**: An invitation letter is sent to each sample address, informing the residents that an interviewer will visit to seek personal interviews. A few days later, interviewers begin visiting sample addresses to seek CAPI interviews. After several weeks of CAPI fieldwork, any sample households that have not yet participated (or have only partially participated with, for example, one household member yet to provide an individual interview) are invited by mail to take part online. For cases where a telephone number is available, a final “mop-up” stage involves telephoning to either encourage online response or obtain a CATI interview.

Group 2 (n = 842 households), **Web-first**: An invitation letter is sent to each sample address, inviting participation in a web survey (see also experiment #2 below). After approximately five weeks, CAPI interviewers begin visiting the addresses of households that have not yet participated (or have only partially participated). For cases where a telephone number is available, a final “mop-up” stage involves telephoning to remind sample members that they can participate online or obtain a CATI interview.

Push-to-web experiment #2: Inviting all household members

Group 1 (n = 422 households), **upfront invites**: The household invitation letter includes a £10 unconditional incentive and a promise of an additional £15 for each person in the household (16+) who completes the questionnaire online within three weeks. Upon completion of the household grid, if there is more than one adult in the household a screen conveys a message along the lines of “We would like to invite <name> to take part in the survey too. They too will receive £15 for doing so by <date>. Please either enter their email address (we will email them their own personal invite) or click here to print an invitation letter with their own unique entry code.”

Group 2 (n = 420 households), **post-grid invites**: The household invitation letter includes a £10 unconditional incentive and a promise of an additional £15 “if you complete the questionnaire online within three weeks”. Upon completion of the household grid, the screen regarding other household members appears, as above.

In both groups, households that have not yet participated receive reminder letters after 7 and 15 working days, again mentioning the £15 conditional incentive and the deadline.

Push-to-web experiment #3: Introducing the CAPI phase

Group 1 (n = 422 households), **explicit CAPI**: The invitation letter states that if they are unable to participate online, there will be an opportunity to be visited by an interviewer instead. The first reminder letter again mentions the interviewer visit option.

Group 2 (n = 420 households), **delayed CAPI**: Neither the invitation letter nor the first reminder letter mentions the interviewer visit option.

For both groups, the second reminder letter announces that an interviewer will call and that each person who takes part, either face-to-face or online, will receive £10.

Results

The experiment provides no evidence (table 10) that the web-first protocol is any less successful than the CAPI-first protocol at achieving household participation, but there is evidence to suggest that the proportion of households responding fully (household grid, household interview and all individual interviews completed) is significantly higher with the CAPI-first protocol (16.8% vs. 12.2% web-first, $P = 0.010$), though the size of the effect is modest.

Neither explaining upfront that all adult household members will be requested to participate nor notifying sample households from the outset that there will be a CAPI follow-up (experiments #2 and #3) had a significant effect on response rates (table 10). Notifying sample households from the outset that there will be a CAPI follow-up (experiment #3) had an effect of borderline significance ($P = 0.08$; table 11) on the mode of participation, with more households participating (at least partly) online if they knew there would be a CAPI

follow-up. This is consistent with the idea that people might be willing to complete a web survey in order to avoid having an interviewer visit their home.

Table 10: Effect on household response rates

	Household response	Full household response
<i>Experiment #1:</i>		
CAPI-first	25.1%	16.8%
Web-first	22.5%	12.2%
<i>P</i>	0.163	0.010
<i>Experiment #2:</i>		
Upfront invites	21.5%	10.9%
Post-grid invites	23.4%	13.6%
<i>P</i>	0.473	0.216
<i>Experiment #3:</i>		
Explicit CAPI	23.1%	11.9%
Delayed CAPI	21.7%	12.6%
<i>P</i>	0.627	0.807

Household response: At least household grid completed;

Full household response: Household grid, household interview and all individual interviews completed;

P-values obtained from design-based Pearson chi-squared tests, taking into account the clustering and stratification of the sample using svy: commands in Stata 15.1.

Table 11: Effect on mode of participation

	Online response
<i>Experiment #3:</i>	
Explicit CAPI	43.0%
Delayed CAPI	30.5%
<i>P</i>	0.084

Online response: Some or all completed interviews completed online; base is participating households (at least grid completed);

P-values obtained from design-based Pearson chi-squared tests, taking into account the clustering and stratification of the sample using svy: commands in Stata 15.1.

d. Mr. Chalk the Teacher: An experimental investigation of children's consistency in reporting their parent's occupations (Vernon Gayle, Roxanne Connelly, Christopher Playford)

Background

There are widespread social, political and economic apprehensions about the extent of social mobility in contemporary societies. A notable concern is the link between social origins in childhood and social destinations in adulthood, which indicates low levels of intergenerational social mobility. In Britain, for example, the Social Mobility Commission⁸ was established to monitor progress towards improving social mobility. There is however only a limited amount large-scale British data appropriate for the detailed empirical study of intergenerational social mobility.

The extent of social mobility is often examined through the assessment of differences between adjacent generations (Goldthorpe, 2016). Historically analyses have tended to focus on fathers and their sons (Erikson, 1984). A standard approach in economics is to measure the differences in incomes (or positions on an income distribution), between parents and their offspring in adulthood (for example see Blanden and Machin, 2008). By contrast sociologists have tended to focus on differences in positions between parents and their children in adulthood, using occupational-based measures of social class (for example see Erikson and Goldthorpe, 1992).

The Study

In Understanding Society (and in the British Household Panel Survey) adult respondents were asked the retrospective questions, *thinking back to when you were 14 years old, was your father working at that time? (W_paju)* and *what job was your father doing at that time? (W_pasoc00)*. These measures are extremely useful in the study of intergenerational

⁸ See <https://www.gov.uk/government/organisations/social-mobility-commission> accessed 16.04.19.

mobility, but have potentially wider uses as background measures of household circumstances in childhood.

At the current time little is known about the accuracy of the recall of these types of measures⁹. In this experiment questions on parents' occupations are asked to the children in the youth sample of the Innovation Panel. The design involves the addition of two questions to the Youth self-completion questionnaire i) What was your father's main job last week? ii) What was your mother's main job last week?¹⁰ The children's data will be compared with the data that are simultaneously collected from their parents in the adult survey.

Theory

We theorise, *a priori*, that there will be a relatively high degree of agreement between the occupations that the children report and the occupations that their parents report. We conjecture that some occupations may be more straightforward, for example traditional and widely comprehended occupations such as butcher, baker, teacher, nurse, police officer, etc. and therefore might be well understood and easier for children to accurately report. By contrast the difference between a dispensing optician (SOC 3216) and an ophthalmic optician (SOC 2214) might be less well understood and lead to misreporting. At first this might not seem unimportant but in practice dispensing opticians are coded to NS-SEC 3 (intermediate occupations) and ophthalmic opticians are coded to NS-SEC 2 (lower managerial, administrative, and professional occupations). Therefore, the misreporting of these occupations would have consequences for measurement of social class and could potentially distort analyses of social mobility.

As SOC codes require additional processing, these data were not available for preliminary release, and analyses were not possible. Future analyses will explore the reports children provide for their parents' work, and any discrepancies with their parents' own reports.

⁹ West et al. (2001) and Vereecken and Vandeghechuchte (2003) are examples of studies that have tested children's reports of their parental occupation but these studies did not collect nationally representative UK data and the data that are now two-decades old.

¹⁰ These questions mirror *Jbsoc00 What was your main job last week?* in the adult survey.

e. Consent to HMRC data linkage experiment (Annette Jäckle, Jonathan Burton, Mick P. Couper, Thomas F. Crossley, Sandra Walzenbach)

This experiment was designed to gain insights into how respondents make the decision whether to consent to their survey responses being linked to administrative records – and to understand why it is that respondents are much less likely to consent if they answer the request online than in a face-to-face interview (see Jäckle et al. 2018; Sakshaug et al. 2017; Thornby et al. 2018).

Data and experimental design

This experiment made use of the mixed mode experiment in the Innovation Panel, where part of the sample are allocated first to CAPI interviewers, and non-respondents are later offered the option to complete the survey online. The rest of the sample are allocated to web-first, with non-respondents followed up by CAPI interviewers (for more information on the mixed mode design, see Jäckle et al. 2019).

All IP11 respondents were asked for consent to link their survey data to HMRC tax records. Respondents were randomly allocated to either an *easy* or *standard* version of the consent question. The standard version was the question that had been used previously in the main Understanding Society sample. The easy wording used shorter sentences and words, no passive sentences, and was visually broken up by using bullet points. In combination this reduced the reading difficulty as measured by the Flesh Reading Ease score implemented in Microsoft Word.¹¹

For CAPI respondents the easy/standard treatment groups were split such that half answered the consent question *early* and half *late* in questionnaire. All web respondents answered the consent question late in the questionnaire.

The consent request was followed by a series of follow-up questions asked in the self-completion section, which in the web version was just a continuation of the questionnaire. These were designed to understand how respondents made their decision, what they had understood about the

¹¹ For question wordings see the IP11 questionnaire at: <https://www.understandingsociety.ac.uk/documentation/innovation-panel/questionnaires>.

consent request, how much they trusted the organisations involved and to what extent they perceived the request as sensitive.

A total of 2,895 respondents gave valid answers to the consent question. However 233 CAPI respondents did not answer the self-completion section containing the follow-up questions and are therefore dropped from the following analyses. Table 12 shows the remaining number of cases in the six experimental conditions.

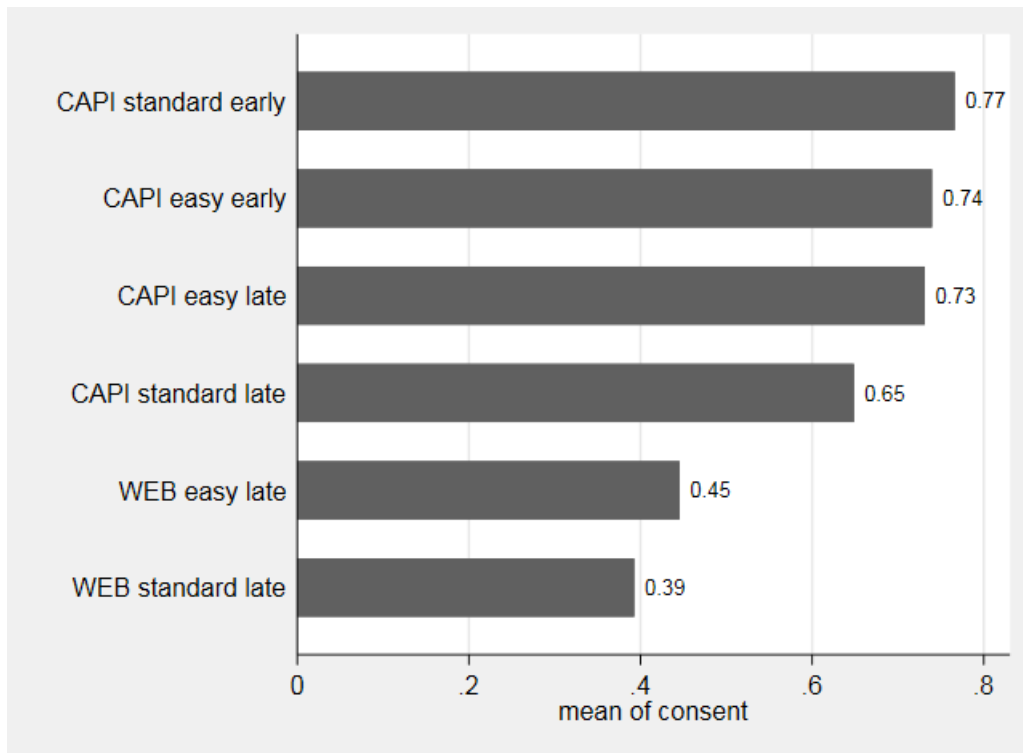
Table 12. Number of cases in experimental conditions

Experimental condition	CAPI	Web	Total
Easy late	320	642	962
Standard late	333	657	990
Easy early	338	0	338
Standard early	372	0	372
Total	1,363	1,299	2,662

Results

A first look at the average consent rate across experimental conditions reveals large mode effects (Figure 5). In line with previous studies (Jäckle et al. 2018; Sakshaug et al. 2017; Thornby et al. 2018), CAPI respondents are significantly more likely to consent (65-77%) than web respondents (39-45%). Within the CAPI group, respondents who received the consent request in the standard wording late in the questionnaire are most reluctant to give consent (65%), while the other groups look rather similar (73-77%).

Figure 5. Average consent rates by experimental condition



The combination of a complicatedly worded request that is asked late in the questionnaire (when respondents might already be tired) has a detrimental effect on consent rates – a difference that is significant compared to all other CAPI groups. Locating the consent question early in the questionnaire raises consent rates in the CAPI standard wording group by 12 percentage points, while it hardly changes consent rates in the condition with easy wording.

Our hypothesis is that these differences in consent rates might reflect how thoroughly respondents processed the consent request and to what extent they understood what they were being asked. We indeed find initial indications that differential processing might play a role.

Column 1 in Table 13 summarises how understanding of the consent request varied with the experimental treatments. It shows the average number of correct answers respondents gave in response to eight true/false knowledge questions about the data linkage process. Column 2 shows the respondents' subjective understanding on a scale from 1 *I do not understand at all* to 4 *I completely understand* (question wording: “How well do you think you understand what would happen with your data, if you allowed us to link it to records held by HMRC?”). The third

column gives some indication about how thoroughly respondents processed the consent request. Respondents were asked “How did you decide whether to say ‘yes’ or ‘no’ in response to the question about data linkage?” (response categories: "I thought about what would happen if I said ‘yes’ or ‘no’, instinct or gut feeling, I said what I usually say when I’m asked for information that is very personal”). The table shows the share of respondents who reported that they considered the consequences of their decision.¹² We assume that respondents who thought about what would happen processed the consent request more systematically, than those who reported answering based on their gut feeling or giving their usual response.

Table 13. Objective and subjective understanding of consent request

	Objective understanding (range: 0-8) (mean)	Subjective understanding (range: 1-4) (mean)	Decision making: “thought about what would happen” (%)
CAPI easy late	5.1	2.9	43.7
CAPI standard late	4.5	2.8	35.7
CAPI easy early	5.0	3.0	39.5
CAPI standard early	4.5	2.8	44.7
Web easy late	4.4	2.4	35.7
Web standard late	3.8	2.4	27.8

Overall, CAPI respondents are not only much more likely to consent than web respondents, they also show higher values in objective and subjective understanding as well as higher shares of respondents who report processing the consent request systematically.

In particular, the web respondents who received the consent question in a standard wording stand out in that they deviate most strongly from all other respondents. The web group who answered an easily worded consent request show similar patterns as the CAPI standard late group with respect to their objective knowledge scores and share of systematically processing respondents. When compared to the CAPI group that similarly answered an easily worded consent question

¹² These analyses are based on the 85.7% of the sample that could clearly be categorised as systematic or unsystematic decision makers. 5.4% of respondents ticked a combination of both strategies; 9% ticked nothing at all.

late in the questionnaire, however, all three indicators suggest significantly less thorough processing and understanding in the web group. Pairwise comparisons of the experimental conditions that only differ in the difficulty of the wording (and not in mode or location of the consent question), show that easier wording significantly enhances objective understanding in all of them. Looking at the two CAPI groups with easy and difficult wording, respectively, reveals that the location of the question in the questionnaire has no effect on understanding. Surprisingly, however, standard wording early in the questionnaire and easy wording late in the questionnaire seem to perform best in encouraging respondents to process the request systematically. However, the differences compared to the other CAPI groups are not always significant.

Preliminary Conclusions and Next Steps

In line with the results of previous studies, our experiments on consent to HMRC data linkage show large differences in consent rates between modes of interview. Early results suggest that these differences can partly be attributed to differences in how thoroughly respondents process and understand the consent request in these two different environments of survey completion. However, what is good for consent rates does not always seem to be good for processing and understanding.

These preliminary analyses do not yet account for self-selection of respondents into modes of interview, and can therefore not be interpreted as causal effects of the mode on how respondents process consent requests. Further analyses will account for selection into modes and examine how the mode affects processing of the consent request, by using additional information such as paradata on response times and whether respondents looked at additional material explaining the data linkage process.

f. Does Competition over Public Services Decrease Support for Residency Rights of Immigrants? Evidence From the United Kingdom (Nicole Martin, Catherine De Vries)

Description of Experiment 1

We prompted adult respondents in the Innovation Panel to evaluate two fictional applicants for permanent residency with the following text: “The government is currently revising the criteria

for eligibility of those applying for permanent residency in the United Kingdom. Please read the descriptions of two applicants carefully. Then, please indicate which of the two applicants you personally think should be granted a permanent right of residence in the United Kingdom.” The two applicants’ characteristics were presented side-by-side in a table, and respondents were asked to indicate their preference with the following question: “If you had to choose between both, which of these two applicants should be given the right to remain in the United Kingdom?”

Description of Experiment 2

In a next step, we explore whether the local context plays a role in respondents’ support for migrants receiving permanent residency. To explore the possible effect of local context, we asked respondents to evaluate an application for the right to remain with the same characteristics, but varied the town in which the applicant lived. Half the respondents were told that the applicant lived in the same town as they resided in, and half were told that they lived in Glasgow. The question wording was as follows: “A male born in Poland who lives in a private rental property [Glasgow/ name of town where the respondent lives] has put in an application for the right to remain in the United Kingdom indefinitely. He has a pre-existing health condition, is married and has three children in state school.”

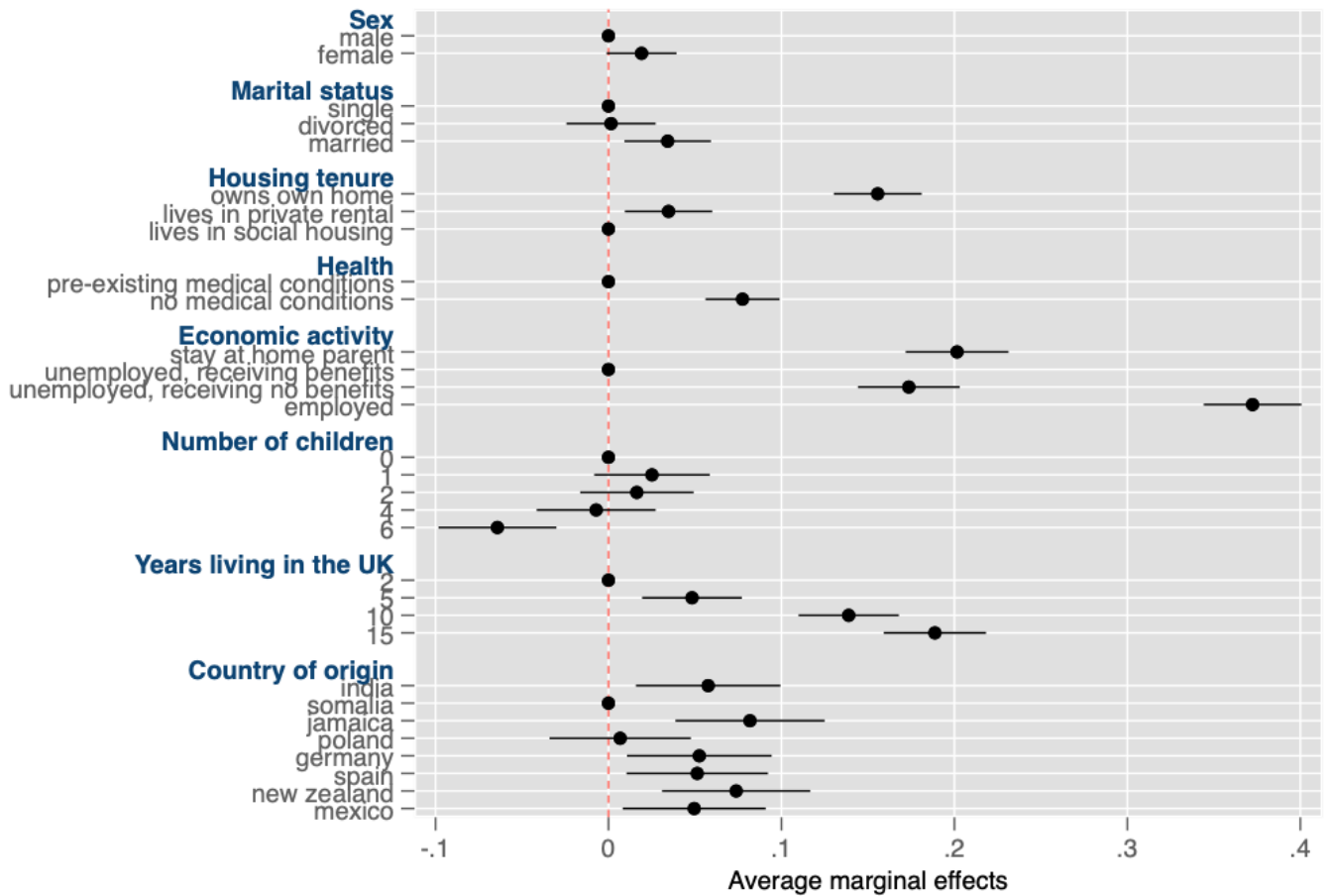
Respondents that received the question” On a scale from 1 to 10, where 1 indicates that the government should not grant the person the right to remain in the United Kingdom and 10 indicates that the applicant should be granted the right to remain in the United Kingdom, how would you rate the applicant?”

Results of Experiment 1

Figure 6 presents the average marginal effects of different migrant characteristics on the probability of being the preferred applicant for permanent residence. Characteristics associated with a heavier reliance on the welfare state appear to count against applicants in our respondents’ minds. The largest gap of 37 percentage points is between an employed applicant and an applicant who is unemployed and in receipt of unemployment benefits. When we specify that the applicant who is unemployed is not in receipt of benefits however, the gap is reduced by half (20 points). A stay-at-home parent is also less preferred than an employed applicant, by 20 points. Other characteristics associated with greater economic self-sufficiency are also important; applicants in social housing were 16 points less likely to be chosen than those who owned their

own home, and 3 points less than those in private renting. It is interesting that the biggest gap is between home owners and both kinds of renters however (13 and 16. points), rather than between those in the public and private sectors of rented accommodation. Applicants with a long-term health condition were 8 points less popular than those in good health. The effect of children on probability of being the preferred applicant appears to take a non-linear form; applicants with only one child and those with two are slightly more preferred, although this difference is not statistically significant. The coefficient for applicants with four children is negative (though still statistically insignificant). It is only when applicants have six children that they are markedly less preferred than those with none, one, two and four. Other characteristics remain important, however. The characteristic of applicants which shows the second biggest difference is years already resident in the United Kingdom; applicants who have lived in the United Kingdom for 10 or 15 years are strongly preferred (14 and 19 points respectively) to those who have lived in the United Kingdom for only 2 years. This suggests that the British public do regard evidence of longer commitment and residence as conditions which should be taken into account in citizenship decisions. The country of origin effects are small in comparison, but applicants from Commonwealth countries (Jamaica, India and New Zealand) are preferred over applicants from Somalia and Poland. Applicants from an older EU member state – Germany – are not significantly different in their rates of being the preferred applicant to those from India or New Zealand however, suggesting that hostility towards immigrants from Central and Eastern Europe extends to those looking to settle in the United Kingdom, as well as shorter-term immigrants.

Figure 6: Average Marginal Effects of immigrant characteristics on likelihood of being granted permanent residence



Results of Experiment 2

Turning to the effect of the local context, we find no significant difference in the evaluation of the applicant for permanent residency based on the location of his residence. The mean score for the applicant who lived in Glasgow was 6.3, compared to 6.2 for the applicant in the respondents home town. This analysis was limited to respondents in England in order to avoid any positive bias towards the applicant in Scotland on the basis of Scottish national identity. In future versions

of this paper we will use the same geocoded data to test whether there is a preference for applicants who live further away in areas with worse public service provision. For this reason, the applicant was described as having characteristics that are associated with greater need for public services – having a health condition and three children.

g. Spending Study 2: Mode of invitation experiment (Annette Jäckle, Alexander Wenz, Jonathan Burton, Mick P. Couper)

All IP11 respondents were invited to participate in a follow-up study to measure household spending, known as Spending Study 2. Participants were asked to download an app onto their Apple smartphone or Android smartphone or tablet, and to use the app for 31 days to report all spending. Within the app, participants were asked to record all direct debits and standing orders that would come out of their accounts during this period. In addition they were asked to record all daily payments, by selecting a payment category and reporting its value. Respondents were offered incentives worth £1 for completing the direct debit/standing order section, £0.50 for every day on which they used the app (including to report days on which they spent no money), a £10 bonus if they used the app every day throughout the month, and £3 for completing a short debrief questionnaire at the end of the study. Incentives were sent to participants by post at the end of the study, in the form of Love2shop gift vouchers.

Experimental design

This study was a follow up to a previous study (Spending Study 1) that had been implemented on the Innovation Panel sample between waves 9 and 10 (see Jäckle et al. 2018). This second study included an experiment that aimed to test ways of increasing participation in app-based data collection: in one half of households respondents were invited to Spending Study 2 within the IP11 interview; in the other half of households respondents were sent a letter inviting them to the study by post, after completing their IP11 interview. The allocation to the invitation treatment was stratified by allocation to the mode of interview for IP11.

The analysis excludes 258 sample members from a reserve refreshment sample issued at IP11, who were invited to Spending Study 2 but for whom the logins were erroneously not activated. The analysis also excludes a single respondent who completed the IP11 interview by telephone

and 90 respondents in households with a low predicted propensity of completing the IP interview online, who were all allocated to CAPI-first rather than being randomly allocated to mode of interview.

Results

Table 14 shows the percentage of sample members who used the app at least once to report a purchase, by mode of IP11 interview and by mode of invitation to Spending Study 2. The results suggest that web respondents were more likely to participate in the Spending Study than CAPI respondents (19.6% versus 15.1%, $P=0.013$) and that inviting respondents to the app study within the IP11 interview achieved higher participation rates than sending an invitation letter by post (22.6% compared to 12.4%, $P<0.001$).

Table 14: Participation rates by mode of interview and invitation

	IP11 respondents		Min. 1 purchases in app		P-values
	N	N	%		
IP11 mode: CAPI	1,261	191	15.1	.	
IP11 mode: Web	1,286	252	19.6	0.013	
Invitation: In-Interview	1,253	283	22.6	.	
Invitation: Post	1,294	160	12.4	0.000	

Notes: P-values from χ^2 tests accounting for complex sample design of IP.

Further analyses reported in Jäckle et al. (2019b) show that in fact inviting respondents to the app study within the IP11 interview only helps if respondents complete the IP11 interview with an interviewer. In that case, participation rates were more than tripled compared to inviting respondents to the study by post. In contrast, if respondents completed the IP11 interview as a self-completion survey online, then the mode of invitation to Spending Study 2 had no effect on whether or not respondents participated. The analyses reported in Jäckle et al. (2019b) also show that although inviting respondents to the Spending Study within the IP interview increased participation, it did not alter the composition of the participant sample in terms of socio-demographic characteristics (gender, age, education, whether the respondent is in work), mobile device usage (how intensely they use their smartphone), financial behaviours (how frequently

they check their bank balance, whether they keep a budget), and financial outcomes (total spending in the last week).

h. Evaluating opt-out consent for linkage to the electoral register (Nicole Martin, Maria Sobolewska)

Voter registration and turnout are key phenomenon the study of electoral behaviour. Accurately measuring these depends on linking survey responses to the electoral register. We ran an experiment in IP10 and IP11 to test out some different aspects of how people are asked for their consent to link to the electoral register. This experiment aimed to assess the extent to which opt-out consent is acceptable to respondents using a two-wave design. The first wave in IP10 found that (i) far fewer people opt-out of consent than refuse if asked outright, and (ii) there is no difference in consent rates between questions that give different reasons for requesting linkage. The second wave of the experiment here considers respondents' own evaluations of whether opt-out consent is acceptable.

Experimental design

Respondents were randomised at the household level in IP10 into 4 conditions which vary on (i) the motivation given for data linkage, and (ii) opt-in or opt-out consent . Those in the opt-out condition were asked a number of questions to ascertain their opinion of opt-out consent in this case. We first asked respondents if they remembered receiving the mailing which contained the information about opt-out consent.

Earlier this year we sent you a letter containing a copy of Insights 2017, containing information about some of the research that has been carried out recently using the study.

Do you remember receiving the information about linking your electoral registration in this letter?

If the respondent reported that they had received this letter, we asked them further questions about it. We asked first if they believed that they were aware of the contents of the letter.

We would like to find out if the information we gave you about this linkage was suitable.

Based on the information provided in this letter, do you think you consented to link your electoral registration to your survey responses?

We then explained the content of the letter and asked whether respondents felt that this was acceptable. If they indicated that they did not, then we provided an open response to explain why. Finally, we asked if they thought that the utility of the linkage had been adequately explained in the letter.

The letter asked that you return a form by Freepost in order to decline linking your electoral registration to your survey responses. If we did not receive this form, it was taken that you consented to link your electoral registration to your survey responses. Do you think this is an acceptable way to obtain your consent for this linkage?

Why do you think it is unacceptable?

Was the usefulness of linking your electoral registration to your survey responses adequately explained in the letter?

We also asked respondents directly if they consented for their details to be linked to the electoral register, so as to have a comparison with the direct opt-in approach.

We want to make sure we accurately record your willingness to consent to this linkage. We would like to link the answers you have given in this survey to other information about the proportion of people round here who voted. To do this we would like your permission to match your name and address to information held on the electoral register. This matching will only be done with information that is already publicly available, and will NOT include any information about who you voted for.

Would you be willing for us to add administrative data from the electoral register to the answers you have given us today?

Results

Of the 1,065 adults in the opt-out condition who took the full adult questionnaire in IP11, 37% reported that they remembered receiving the information about linkage. 60% said they did not, and 4% said they did not know or refused to answer the question. Of the 389 who recalled the information, 48% reported that they thought that they had consented to the electoral linkage, with 40% saying that they had not. This 40% is 155 people, but given that only 78 opt outs were received, it is clear that some people believed that they had not consented when the opt-out

consent mechanism would say that they had. 11% did not know, and 1% refused to answer the question. 74% of the 389 respondents reporting having received the letter also felt that the usefulness of linkage to the electoral register was adequately explained in the letter.

These results cast some worries for proponents of opt-out consent; 6 out of 10 respondents did not remember the letter where opt-out consent was explained, and more people believe that they did not consent to this linkage than actually opted out. This suggests that for types of data linkage where informed consent is important, opt-out consent will fail to properly inform or consult with respondents.

After asking about their recollections, all respondents in the opt-out consent condition had the consent procedure explained to them, and their opinion as to its acceptability was sought. The 1,065 adults were evenly split; 46% said that this was acceptable, and 46% said that it was not acceptable. 8% did not know, and 1% refused to answer. Those who remembered receiving the letter were much more likely to say that it was acceptable – 61% felt it was, compared to 37% of those who did not recall.

Finally, respondents were asked for their consent to link to the electoral register using the same wording as in IP10 and the British Election Study. 64% consented, with 34% refusing consent, and 2% refused to answer the question or said that they did not know. This consent rate is lower than the 75% consent rate that this question achieved in IP10. This suggests that opt-out consent has the possibility to damage respondent's trust and should therefore be used sparingly.

References

- Blanden, J. and Machin, S. (2008). Up and down the generational income ladder in Britain: Past changes and future prospects. *National Institute Economic Review*, 205, 101-116.
- Converse, J. M. (1976). Predicting No Opinion in the Polls. *Public Opinion Quarterly*, 40, 515-530.
- Dillman, D.A. (2017) The promise and challenge of pushing respondents to the web in mixed-mode surveys, *Survey Methodology* 43(1): 3-30.
- Erikson, R. (1984). Social class of men, women and families. *Sociology*, 18, 500-514.
- Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*, Oxford University Press, USA.

- Goldthorpe, J. (2016). Social class mobility in modern Britain: changing structure, constant process. *Journal of the British Academy*, 4, 89-111.
- Herdman, M., Gudex, C., Lloyd, A. , Janssen, M., Kind, P. , Parkin, D., Bonsel, G. and Badia, X.(2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 20:1727-1736.
- Hernández-Alava, M. and Pudney, S.E. (2017). Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis, *Journal of Health Economics* **55**, 139-152.
- Hernández-Alava, M., Wailoo, A., Grimm, S., Pudney, S. E., Gomes, M., Sadique, Z., Meads, D., O'Dwyer, J., Barton, G. and Irvine, L. (2018). EQ-5D-5L versus 3L: the impact on cost-effectiveness in the UK, *Value in Health* **21**, 49-56.
- Hernández-Alava, M., Wailoo, A. and Pudney, S.E. (2019). Mapping between EQ-5D-3L and EQ-5D-5L. A survey experiment on the validity of multi-instrument data. University of Sheffield: HEDS Working Paper (in preparation).
- Jäckle, A., T. Al Baghal, J. Burton, O. Kaminska, and P. Lynn (2019) Understanding Society: The UK Household Longitudinal Study Innovation Panel, Waves 1-11, User Manual. Colchester: University of Essex.
- Jäckle, A., Beninger, K., Burton, J., & Couper, M. (2018) Understanding data linkage consent in longitudinal surveys. *Understanding Society Working Paper 2018-07*. Colchester: University of Essex.
<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2018-07.pdf>
- Jäckle, A., Burton, J., Wenz, A., and Read, B. (2018) [Understanding Society The UK Household Longitudinal Study: Spending Study 1 User Guide](#), Colchester: University of Essex.
- Jäckle, A., Jonathan Burton, A. Wenz, B. Read, T. Hanson, and D. Xu (2019a) Understanding Society: The UK Household Longitudinal Study Spending Study 2 User Guide. Colchester: University of Essex.
- Jäckle, A., Wenz, A., Burton, J., and Couper, M.P. (2019b) Increasing participation in a mobile app study: the effects of a sequential mixed-mode design and in-interview invitation. *Understanding Society Working Paper 2019-04*. Colchester: University of Essex.
- Sakshaug, J. W, Hülle, S., Schmucker, A., & Liebig, S. (2017). "Exploring the effects of interviewer- and self-administered survey modes on record linkage consent rates and bias." *Survey Research Methods* 11(2): 171-88.
- Thornby, M., Calderwood L., Kotecha M., Beninger K. & Gaia A. (2018). Collecting Multiple Data Linkage Consents in a Mixed-mode Survey: Evidence from a large-scale longitudinal study in the UK Survey Methods: Insights from the Field. Retrieved from <https://surveyinsights.org/?p=9734>

- Vereecken, C. and Vandegehuchte, A. (2003). Measurement of parental occupation: Agreement between parents and their children. *Arch Public Health*, 61, 141-9.
- West, P., Sweeting, H. and Speed, E. 2001. We really do know what you do: a comparison of reports from 11 year olds and their parents in respect of parental economic activity and occupation. *Sociology*, 35, 539-559.