Understanding Society
Working Paper Series

No. 2018 – 07

June 2018

Understanding data linkage consent in longitudinal surveys

Annette Jäckle[1]
Kelsey Beninger[2]
Jonathan Burton[1]
Mick P. Couper[3]

[1]University of Essex
[2]Kantar Public
[3]University of Michigan

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Understanding Society
THE UK HOUSEHOLD LONGITUDINAL STUDY

Non-technical summary

As new sources of data emerge, such as data generated by social media or private companies, there is increasing interest in supplementing data collected in probability sample surveys with such data. In many countries survey data can only be linked to administrative or other process-generated data if survey respondents give informed consent to the linkage. In this paper we focus on how respondents in longitudinal surveys make the decision whether or not to consent, and how this decision is influenced by the mode of data collection.

We draw on both quantitative and qualitative analyses. We first analyse consent questions in *Understanding Society*: the UK Household Longitudinal Study and the Innovation Panel, asked over time and in different modes. Second we report on results from qualitative in-depth interviews carried out in spring 2017 with 25 members of the Innovation Panel sample.

The quantitative analyses show that respondents are up to 30 percentage points less likely to give consent when they complete the survey online than when they complete a face-to-face interview. This appears to be due to the mode affecting willingness to consent, not due to the type of people responding online. In addition, nearly 50% of respondents who do not give consent do consent when asked again in a later wave. This suggests that the consent decision is not a strong or fixed decision, but can likely be influenced.

The qualitative interviews show that there is a lot of misunderstanding by participants of what the consent question is asking: about how personal data would be used, by whom and for what purposes. Factors influencing the participant's decision whether or not to consent include personality traits (such as openness versus suspiciousness), attitudes towards data sharing, assessment of the benefits and risks of granting access to personal data, and trust in the survey organisation. Participants gave several reasons why they would be less willing to give consent in an online survey: the difficulty of reading a lot of content on a computer screen and concern about privacy and security of personal information shared online. The possibility of asking the interviewer clarification questions and obtaining reassurance were further reasons why participants thought they would be more likely to give consent in the face-to-face setting. The reactions of participants who had given inconsistent responses to consent questions at different points in time highlighted the role of subconscious, environmental and social factors in driving decisions to consent. Participants were generally unable to recall whether they had consented in the past, and were sometimes surprised about their previous responses.

This paper ends with a discussion of practical implications of the quantitative and qualitative findings, and a discussion of further research needs on methods to increase informed consent.

Understanding data linkage consent in longitudinal surveys

Annette Jäckle (University of Essex, aejack@essex.ac.uk)
Kelsey Beninger (Kantar Public, kelsey.beninger@kantarpublic.com)
Jonathan Burton (University of Essex, jburton@essex.ac.uk)
Mick P. Couper (University of Michigan, MCouper@umich.edu)

Abstract

We explore the cognitive process by which respondents decide whether or not to consent to having their survey data linked to administrative records. Using data from the *Understanding Society* household panel study we document the extent of inconsistencies in respondents' consent decisions, between data domains and over time, and the impact of the mode of data collection on consent. To understand the reasons for inconsistencies and why the mode affects consent, we report on qualitative in-depth interviews that examine how respondents understand the consent request and which factors influence their decision.

Keywords: record linkage, mode of data collection

JEL classification: C81, C83

Acknowledgements*:* A revised version of this article will be published as a chapter in the book 'Advances in Longitudinal Survey Methodology' (editor P. Lynn), to be published by Wiley in 2019. Readers wishing to cite this article should cite the book version, not this Working Paper. This research was funded by the UK Economic and Social Research Council grant for *Understanding Society* [ES/N00812X/1] and by the Nuffield Foundation [OSP/43279], but the views expressed are those of the authors and not necessarily those of the Foundation.

Corresponding author: Annette Jäckle, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Essex, CO4 3SQ, aejack@essex.ac.uk

1. Introduction

Linking data from longitudinal surveys to administrative records (whether held by government or private entities) is an increasingly attractive option for several reasons. First, the rising costs of survey data collection and declining response rates require researchers to look elsewhere to supplement or replace survey data. Second, the increasing demand for more detailed and timely data raises concerns about respondent burden. Third, the increasing availability of digital data from a variety of sources, along with the development of improved tools for data linkage and protections against disclosure opens up new opportunities. The various ways in which administrative data can be used, whether to replace or enhance surveys, are detailed elsewhere (e.g., Calderwood and Lessof, 2009; Gates, 2011; Groves and Harris-Kojetin, 2017, p. 33-34). We take as given that the demand for administrative data linkage in surveys is likely to increase.

Administrative data linkage is not without challenges. A key one is that of obtaining informed consent for respondents to link their survey data to administrative records. While some statistical agencies (e.g., Statistics Canada; U.S. Census Bureau; more recently, with the passage of the Digital Economy Act, the UK Office for National Statistics) are taking the view that explicit consent does not need to be solicited for such linkages, such an option is generally not available to non-government research organisations. In fact, the newly introduced European General Data Protection Regulation focuses on the need to obtain explicit informed consent. Further, regardless of the legal framework for administrative data linkage, ethical issues remain (Lessof, 2009). This means that addressing the challenges of gaining consent, and ensuring that the consent is informed, remains a focus for methodological research.

Failure to obtain consent from respondents has several potential consequences. First, the possibility of non-consent bias may arise if those who consent are different from those who do not consent. Second, even if bias is not present, low consent rates may mean smaller effective samples for those using the administrative data. Third, if administrative data are used to replace survey questions, non-consenters may need to be asked additional questions, increasing respondent burden. Finally, the act of requesting consent to administrative data may be viewed as intrusive or threatening to respondents, potentially leading to panel attrition.

Much of the work on informed consent for administrative data linkages has focused on exploring correlates of consent at both the respondent and interviewer level and examining non-consent bias (e.g., Al Baghal, 2016; Fulton, 2012; Korbmacher and Schroeder, 2013; Knies and Burton, 2014; Mostafa, 2014; Mostafa and Wiggins, 2018; Sakshaug et al., 2012; Sakshaug, Tutz, and Kreuter, 2013; Sala, Burton, and Knies, 2012). Most of these studies involve secondary analysis of existing data, often focusing on a single survey (Knies and Burton, 2014, is an exception). These studies find considerable variation between respondents and between interviewers in consent rates, but the findings with regard to the effect of specific characteristics of respondents and interviewer on consent are inconsistent across studies.

Experimental studies examining how the consent question is worded (Das and Couper, 2014; Pascale, 2011; Sakshaug, Tutz, and Kreuter, 2013), or whether it is asked in an earlier or later wave of a panel survey (Eisnecker and Kroh, 2016; Sala, Knies, and Burton, 2014) have generally found few consistent effects. Asking for consent after a module of questions related to the content of the data to be linked increased consent compared to asking at the end of the questionnaire (Sala, Knies, and

Burton, 2014). Wording the consent request to emphasise that not linking will reduce the value of the respondent's survey data (loss framing) increased consent compared to gain framing which emphasises the value of linkage (Kreuter, Sakshaug, and Tourangeau, 2016).

There is a relative paucity of research exploring the reasons for non-consent and identifying strategies for addressing these reasons. Similarly, there is little work focusing on how informed respondents are – whether objectively or subjectively measured – about their decision to consent or not (for exceptions, see Das and Couper, 2014; Thornby et al., 2017).

There is some evidence that consent rates are declining over time in longitudinal studies, which raises further concerns about potential bias. For example, among persons newly-eligible to be asked consent to Medicare linkage in the U.S. Panel Study of Income Dynamics (PSID), consent rates have declined over time, from 60% in 2005, to 41% in 2007, 48% in 2009, 29% in 2011, 36% in 2013, and 32% in 2015 (Freedman, McGonagle, and Andreski, 2014; Katherine McGonagle, personal communication). Similarly, in the U.S. Health and Retirement Study (HRS), 67% of those interviewed for the first time face-to-face in 2004 consented to Social Security Administration linkage, while for those interviewed for the first time in 2010 (new cohorts are introduced every 6 years), the corresponding number was 62%.

There is also evidence from longitudinal surveys that consent varies within respondents over time. It is common practice to re-ask consent at later waves of those who did not consent, but some surveys also repeat the request of all respondents. Mostafa and Wiggins (2018) found that 76% of respondents consented to linkage to children's health records in each of three waves of the UK Millennium Cohort Study (MCS), while only 0.5% did not consent in any of the three waves; the balance (24%) consented in one or two waves. Weir, Faul, and Ofstedal (2014) report that in the HRS, 42% of respondents who had not given consent to linkage with Medicare data gave consent the second time they were asked. Similarly, the PSID, which predominantly uses the telephone mode of data collection, has sought oral consent for those eligible for Medicare (primarily persons 65 and older) since 2005 (Freedman, McGonagle, and Andreski, 2014). Consenting respondents were asked to provide their Medicare number. In 2005, 60% of eligible persons consented and provided a valid number. Those who did not consent or did not provide a valid number were asked again in subsequent waves. Through 2015, a further 16% consented, yielding an overall consent rate of 76% among those eligible in 2005.

As longitudinal surveys continue to respond to pressures to increase efficiency, and as new survey modes are developed, the use of mixed-mode data collection is increasing. This presents additional challenges for consent. Evidence from a pilot study for the Next Steps cohort study showed that the average consent rate across 9 different requests was highest in face-to-face interviews (78%), lower in telephone (71%), and lowest in web (61%, Calderwood, 2016; Thornby et al., 2017). Generally not much is known about why consent rates are lower in web, or how best to obtain informed consent from web respondents.

There is considerable variation in how consent is requested or obtained in longitudinal surveys. Knies and Burton (2014) document some of that variation in UK health surveys. Some surveys require a signature to document consent, or provision of an identification number (e.g., Social Security Number) to confirm consent and facilitate linkage, despite evidence that this may increase respondent concerns about disclosure risk (see Singer, 1978, 2003) and lower consent rates (Dahlhamer and Cox, 2007; Bates, 2005). Sometimes the consent is for retrospective data linkage (e.g., data existing at the time of the interview); other times it is prospective (e.g., administrative data generated in the future). There is also variation in whether respondents are asked to confirm their consent and, if so, at what interval (asked just once or repeated requests); whether consent is opt-out (increasingly used by government agencies with authority to link survey and administrative data) or opt-in (often required for academic or social research); where in the survey consent is asked (at the end or in the topical section); how the consent request is framed or worded; and how to ask multiple consent questions. This variation may account for some of the inconsistent findings in the literature, but also points to a need for more research on consent to administrative record linkage, especially in the context of longitudinal surveys where multiple consent requests (both within and between waves) are common, and mixed-mode data collection is increasingly the norm.

Given this background, the focus of this chapter is on understanding the process by which respondents decide whether or not to give consent. We use both quantitative analyses of existing data from *Understanding Society* and qualitative work to explore the consent process. The quantitative analyses document the extent to which respondents make consistent decisions, and the extent to which the mode of data collection impacts on this decision, by addressing the following research questions:

- How consistent are respondents about giving consent to data linkage between topics and over time?
- Does consistency over time vary between domains?
- What is the effect of the survey mode on consent? Are differences in consent rates due to selection effects or does mode affect the consent decision?

To understand the reasons for inconsistencies in consent decisions, and why the mode of data collection has an effect, the in-depth qualitative interviews address the following research questions:

- How do participants interpret consent questions?
- What do participants think are the implications of giving consent to linkage?
- What influences the participant's decision whether or not to give consent?
- How does the survey mode influence the decision to consent?
- Why do participants change their consent decision over time?

2. Quantitative research: consistency of consent and effect of mode of data collection

2.1 Data and methods

To examine the consistency of consent over time and across topics, we use the main sample of *Understanding Society* (University of Essex, 2017a). Over the course of the main-stage *Understanding Society* survey, respondents were asked in the CAPI interview for consent to link administrative data in multiple domains to their survey responses. Table 1 shows the waves of data collection in which consent to link was asked of adults (aged 16+) in the individual interviews, the unweighted consent rate, and the number of respondents who were asked for consent at each request. This excludes cases where consents were asked of the responsible adult on behalf of children, and consent to link to energy consumption data which was asked in the household questionnaire. For more information about data linkage on *Understanding Society*, see https://www.understandingsociety.ac.uk/about/data-linkage. For details on the wording of the consent questions,[1] and the participant materials used,[2] refer to the study website.

In general, at the first time of asking for consent to link to data on a specific domain, all eligible adults were asked. Eligibility for linkage to education data was based on age and country in which the person went to school; for the other linkage requests all adults were eligible. In most subsequent requests for the same domain, only those who had not given consent previously were asked. This included those who had been asked but did not consent, new entrants to the study, and those who had turned 16 and become eligible for the adult interviews since the previous request. At Wave 4, however, those who had previously consented to education and health data linkage were asked to confirm their consent during the interview.

For the first seven waves of *Understanding Society*, the primary mode of data collection was face-to-face, using CAPI (Computer Assisted Personal Interviews). At Wave 7, adults in households where no-one had responded at Wave 6 were invited to participate online. Those that did not complete their interview online after the first few weeks were issued to CAPI interviewers. This design is referred to as a "web-first" design (Jäckle, Lynn, and Burton, 2015). At Wave 8, the group of sample members invited to participate online increased, and included a substantial proportion of those in households that had responded in the previous wave. Altogether, around 40% of households were issued "web-first". Consent questions were asked in both the online and CAPI interviews. For Wave 8 we currently only have data from the first year of the two-year data collection cycle. Therefore only some of the analyses include Wave 8 data, as documented below.

To examine the effect of survey mode on consent, we use the ninth wave of the *Understanding Society* Innovation Panel (IP9) (University of Essex, 2017b). Since Wave 5 of the Innovation Panel (IP5) we have implemented a mixed-mode design in which two-thirds of the sample was issued to web first, with non-respondents followed up by CAPI interviewers. The remaining one-third of sample members were issued directly to interviewers (CAPI-first). Towards the end of the fieldwork period, during the re-issue stage, non-respondents in the CAPI-first group were invited to complete their interview online. At the end of fieldwork, all non-respondents were eligible for a telephone 'mop-up', at which point interviewers would attempt to conduct a CATI interview. The initial random allocation was made at IP5 and sample members retained their mode of issue at subsequent waves.

At IP9 all adult respondents were asked for their consent to link their survey responses to administrative data that the Financial Conduct Authority (FCA) compiles from Credit Rating Agencies (see Jäckle et al., in press). The FCA consent question was asked on the CAPI and online versions of the questionnaire, but not on the telephone.

All significance tests account for the clustering and stratification of the sample. The wording of all consent questions used in the analysis is documented in the Appendix.

---

[1] https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires

[2] https://www.understandingsociety.ac.uk/documentation/mainstage/fieldwork-documents

Table 1: Consent questions in the main stage of *Understanding Society*

Domain	Data-holder	Wave and Eligible sample members	Consent rate	Observations
State benefits	Department for Work and Pensions (DWP)	4 – all 7 – all without valid consent	63.9% 68.7% **	41,988 19,683
Employment/self-employment	Her Majesty's Revenue and Customs (HMRC)	5 – all 8 – all without valid consent	62.7% 45.3% *	39,905 12,075
Vehicles	Driver and Vehicle Licensing Agency (DVLA)	5 – all with driving licence, access to car/van, has at least one vehicle registered in the UK	79.7%	27,382
Education (National Pupil Database, Early Years Census)	Department for Education (DfE)	1 – adults born after 1981 4 – adults born after 1983 who are new entrants or have not given consent, adults who had previously consented (confirm); 5 – 16-year olds 6 – 16-year olds 7 – all without valid consent	77.8% 69.9% 52.1% 72.5% 83.6%	6,091 5,790 652 473 5,7,173
Education – Individual Learning Records	Skills Funding Agency/Department for Business, Innovations, Skills	7 – 16-year olds, OR (born after 1978 AND new entrant since Wave 4/no consent given)	84.7%	8,551
Higher Education	Higher Education Statistics Agency (HESA)	5 – all who completed higher education in the UK in 1995 or later	69.7%	864
NHS Data (Hospital Episode Statistics)	National Health Service (NHS)	1 – all 4 – all (previous consenters asked to confirm, previous non-consenters re-asked) 5 – 16-year olds 6 – 16-year olds	68.1% 50.6% 49.4% 65.4%	47,378 22,454 652 480
NHS Central Register	National Health Service (NHS)	1 – all 4 – all without valid consent 5 – 16-year olds 6 – 16-year olds	67.2% 50.6% 49.4% 65.4%	47,356 22,454 652 480

* Wave 8 data are from year 1 of wave 8 only; year 2 data are not available at the time of writing.

** Due to an error with the feed-forward data which controlled routing in the questionnaire, for a three-month period, all issued sample members were eligible to be asked for consent to the DWP linkage at Wave 7, rather than just those who had not previously consented. These extra cases are included in the analyses.

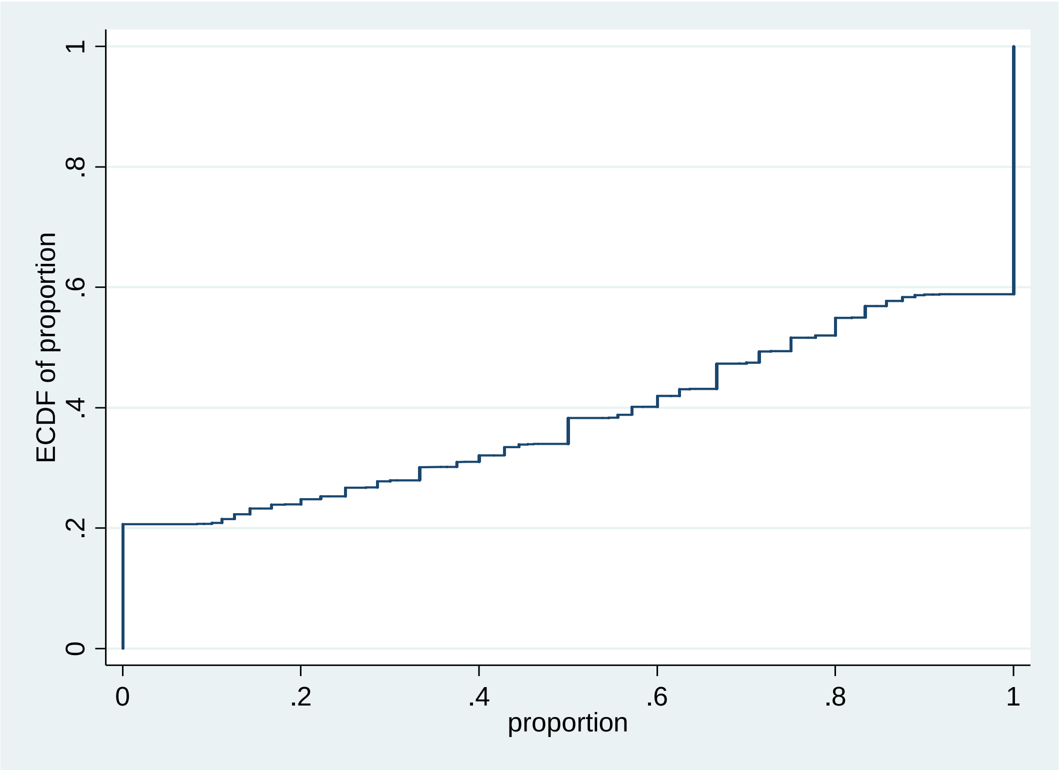
2.2 Results

How consistent are respondents about giving consent to data linkage between topics?

This analysis uses information on all the adult-level consent requests on the main survey. Across all domains of consent and all requests for consent we have a consent rate of 66.1%. Looking at the consent rate for the first time each domain was asked, we see there is variation across domains from 84.7% for the Individual Learning Record (education) to 62.7% for HMRC tax records (Table 1).

To examine the consistency of consenting, we calculated the proportion of consent questions asked of the respondent to which they gave consent. The number of consent requests made to respondents depends on their eligibility, which varies by domain and by the number of waves a respondent has participated in. The figure below shows a cumulative distribution of this indicator: 15.0% of respondents never gave consent to linking their data, whilst almost one-half (48.0%) gave consent each time they were asked. Just over one in three respondents sometimes gave consent and sometimes withheld their consent.

Figure 1: Cumulative distribution of the proportion of consent requests granted



The consistency of consenting documented in Figure 1 may be confounded by the number of consent questions asked of each respondent. It may be, for example, that those who consented to none of the requests were in fact only asked one consent question. Table 2 however suggests that this is not the case. The average number of consents asked of any respondent is 5.9, with the maximum being 15. However, only 4.8% of respondents were asked for consent only once, whilst one-fifth of respondents were asked twice (20.8%). Table 2 shows the average consent rate by the number of consent requests. The consent rate is highest for those respondents who have been asked 4-6 times. The rate is lowest for those who have had 7 or more requests; this may reflect the fact that those who do not consent to linkage to a particular domain of administrative data are re-asked the next wave this is carried.

Table 2: Average consent rate by the number of consent requests

Number of consent requests	Mean consent rate (%)	Std. Err.	95% Confidence Int.		n
1	56.7	0.008	55.1	58.3	3,689
2	58.6	0.004	57.8	59.3	16,039
3	58.2	0.004	57.4	59.1	10,840
4	66.1	0.005	65.2	67.0	7,953
5	79.0	0.003	78.5	79.6	11,883
6	65.8	0.003	65.2	66.5	8,032
7+	53.4	0.002	52.9	53.8	18,831

To formally test how much of the variation in consent rates is at the respondent level, we ran a null multi-level model of the probability of giving consent, with interviewers and respondents as the two levels, and accounting for the clustered sample design. The data structure is such that there are 298,346 consent questions, nested in 77,267 respondents, nested in 1,421 interviewers. The intra-class correlations derived from the null model confirm that there is large variation between respondents in the probability of giving consent (ICC=0.595), and also some variation between interviewers (ICC=0.156).

How consistent are respondents about giving consent to data linkage over time?

We examine consistency of consent over time using measures that examine whether people revoke consent after they have given it, whether they consent to any other request after they have revoked a specific consent, and whether they make the same consent choice at consecutive requests.

First, we examine whether people who give consent, later revoke that consent. Periodically, those who have given consent are sent a letter to remind them of their consent, and to give them a chance to revoke that consent by completing a form and sending it back. In addition, sample members can withdraw their consent at any time by writing to the managers of the study. Revoking consent is however rare: overall, only 3,592 consents have been revoked (1.8% of consents).

Second, we examine whether people who revoke consent at one point in time, then give consent to a different data linkage request at a later time. Surprisingly we find that over three-quarters of those who revoke a consent then agree to consent the next time they are asked (78.6%, n=3,483). Note that those who revoke consent to a particular domain are not re-asked for consent to the same domain the next time this is asked.

Third, we look at whether respondents who do not give consent at one request, then give consent the next time they are asked *any* consent question. Less than one-third of those who do not consent at one request, then grant consent the next time they are asked any consent question (26.9%, n=84,298). This compares to 85.5% of those who do give consent to one request, who then give consent at the next request (n=156,164).

Fourth, we examine whether respondents who did not give consent at one request, then give consent the next time that *same* consent question is asked. We restrict the analysis to those consent requests which have been asked more than once (DWP, Education, NHS Data, NHS-Register). Usually only respondents who did not give consent the first time are asked again. If people withheld their consent for the same domain of data twice, or if they gave consent but later revoked it, they were not asked for consent for that domain again. For part of the first year of Wave 7, however, everyone was asked for consent to education and DWP linkage again, regardless of whether they had previously given consent or not. This was due to an error in the sample data that was issued for the July-September samples in the first year of Wave 7. Overall, we find that just under half of those who did not consent at one request, do give consent when they are asked the same request in a later interview (46.6%, n=31,802). This compares to three-quarters of those who have previously given consent consenting again (72.2%, n=6,412).

Does consistency over time vary between domains?

To examine whether the consistency of consent decisions over time varies between domains of data linkage, we again use those consent requests which have been asked more than once (DWP, Education, NHS). Table 3 shows that consistency does vary between data domains. Among respondents who did not consent previously there are significant differences in the consent rates at *t+1* between domains, ranging from around 39% for health administrative data, to 69.2% for education data (Pearson design-based F(1.71, 5688.87)= 518.60, P < 0.001). Among respondents who did previously consent there is also variation between domains, although to a lesser extent, ranging from around 67% for health-related data to 82.7% for education data (Pearson design-based F(1.75, 2408.91)= 97.23, P < 0.001). Note that only younger adults are eligible to be asked for consent to link to education data, whereas all responding adults are eligible to be re-asked for consent to linkage to the other three domains.

Table 3: Consent given at t+1 by consent status at time t.

Domain	Consent not given	n	Consent given	n
Education	69.2%	1,978	82.7%	1,650
DWP	64.3%	8,209	82.2%	2,873
NHS-Register	39.5%	11,181	67.4%	5,009
NHS Data	39.3%	10,962	66.8%	5,246

Table 4, below, shows the proportion of consents that were revoked by domain of administrative data. Generally, the revocation rate is low, suggesting that once consent has been given, respondents are unlikely to revoke it in the future.

Taken together, this suggests that the decision to consent varies across domains, but also within participants. Those who consent are more likely to consent again, but are not certain to do so – and those who do not consent are less likely to consent in the future, but, again, are not certain to refuse.

Table 4: Proportion of consents revoked by domain.

Domain	Revoke	n of consents revoked
NHS data	2.5%	1,124
NHS Register	2.5%	1,107
Education	1.6%	251
DWP	1.9%	766
DVLA	0.7%	155
HESA	0.8%	5
HMRC	0.7%	183
ILR	0.0%	1

What is the effect of the survey mode on consent? Are differences in consent rates due to selection effects or does mode affect the consent decision?

To look at the effect of survey mode on consent, we use the IP9 data which employs a mixed-mode (web-CAPI) design for two-thirds of the sample. The focus is on consent to link to information held by the Financial Conduct Authority (FCA). In IP5, two-thirds of households were randomly allocated to a sequential web-first design, where non-respondents to the web survey were followed up by face-to-face interviewers. The remaining one-third of households were allocated to CAPI-first. In waves 6 to 9 the allocation of sample members to modes remained unchanged. Table 5 documents the number of interviews in each mode in IP9, by mode to which respondents were allocated. Proxy respondents (63 cases) are dropped as these are not asked for consent to data linkage. In addition, the small number of interviews that were done by telephone are excluded from the analyses (31 cases). As Table 5 shows, among full respondents issued to CAPI-first, 92.1% completed the survey with an interviewer. Among those issued to web-first, 73.4% completed the interview online, while 25.2% completed it with an interviewer.

Table 5: Innovation Panel wave 9, mode of interview by mode of issue (full respondents only)

	CAPI		CATI		Web		Total
Allocation	N	%	N	%	N	%	N
CAPI-first	650	92.1	11	1.6	45	6.4	706
Web-first	370	25.2	20	1.4	1,078	73.4	1,468
Total	1,020	46.9	31	1.4	1,123	51.7	2,174

Although the allocation to modes was carried forward from wave 5, by wave 9 the randomised groups were still comparable. As Table 6 shows there is no difference in IP9 between the CAPI-first and the web-first groups in terms of gender, educational qualifications, sample origin, labour market activity, and mean income. There are differences in terms of age, with web-first sample members being on average nearly 3 years younger than CAPI-first sample members. The last columns in Table 6 however suggest that age is not related to the probability of giving consent. The analyses below therefore make use of the randomisation to mode of interview, to identify the effect of the mode of data collection on consent.

Table 6: Sample composition by issued mode, mode of interview, and consent outcome

	Issued mode			Mode of interview			Consent		
	CAPI-first	Web-first	P-value	CAPI	Web	P-value	No	Yes	P-value
Female	54.4	53.8		54.9	55.21		46.6	53.4	
Male	45.6	46.2	0.6344	45.1	44.79	0.8314	38.5	61.5	0.0000
Degree	24.2	25.2		21.4	27.9		36.6	63.4	
Other higher qualification	13.6	13.1		12.9	14.0		42.2	57.8	
A-level	21.6	22.8		21.5	23.3		45.3	54.8	
GCSE	24.7	25.0		25.5	24.1		44.6	55.4	
Other qualification	7.1	5.6		7.1	4.9		43.9	56.2	
No qualification	9.0	8.4	0.7323	11.6	5.8	0.0000	51.1	48.9	0.0229
Original sample	49.1	49.2		46.6	51.5		44.2	55.9	
IP4 refreshment	25.3	22.2		22.6	23.8		40.5	59.5	
IP7 refreshment	25.7	28.5	0.3944	30.9	24.8	0.0634	42.9	57.1	0.5151
Self-employed	7.1	8.1		7.4	7.8		46.1	53.9	
Employee	45.0	49.3		41.7	52.9		41.7	58.4	
Unemployed	3.0	3.0		3.2	2.8		48.4	51.6	
Retired	30.5	26.1		31.4	25.0		42.0	58.0	
Other	14.4	13.5	0.2475	16.4	11.5	0.0000	44.7	55.3	0.7178
Age (mean)	51.8	49.0	0.0033	52.2	48.3	0.0000	49.9	50.2	0.7841
Gross monthly income (mean)	1812	1836	0.6988	1715	1937	0.0004	1642	1968	0.0001

Notes: P-values from Chi-square tests and tests of means, adjusted for sample design.

Examining the characteristics of respondents by mode in which they completed the wave 9 interviews shows that web and face-to-face respondents are not comparable, and some of the characteristics in which they differ are related to the probability of giving consent (Table 6). Web respondents had higher educational qualifications, were more likely to be in employment, younger, and had higher incomes than face-to-face respondents. In turn, those with higher qualifications and higher income were more likely to give consent. These results confirm the need to disentangle differences in consent rates that are due to different types of people answering in different modes from the causal effect of the mode on consent.

Overall, there was a 56.9% consent rate to link FCA administrative data to the IP9 survey responses. However, as shown in Table 7, the consent rate was 19.0 percentage points higher for those interviewed in-person (67.1%) than those who were surveyed online (48.1%).

The large difference in consent between modes may be related to the positive influence of the interviewer in being able to answer questions the participant may have about consenting, or the tendency – perhaps – for an online respondent to skim over the consent question and answer without giving much thought. That is, the difference may be due to mode differences in how consent is asked. However, the difference may also be due to selection differences into mode.

Table 7 documents the results of three different statistical approaches to estimating the effect of mode on consent, by accounting for differences in selection. In all analyses in Table 7 those individuals who were issued to a face-to-face interviewer but completed their interview online (45 cases) are excluded. For the instrumental variables method this is necessary for the assumptions underlying the method to hold, as explained below. For consistency we maintain the same sample selection criteria for the other analyses.

The first method is to estimate the effect of the 'intention to treat', that is, the consent rate by the randomised mode of issue rather than by the actual mode of interview. For respondents who were issued to CAPI first, the consent rate was 72.3%, for those issued to web-first 50.8%, a difference of 21.5 percentage points. The intention to treat analysis suggests that the difference in consent between modes of interview is not due to selection effects. It does however not provide an estimate of the size of the effect of mode on the consent decision.

The second approach is to control for respondent characteristics which explain the selection into mode (Vannieuwenhuyze, Loosveldt, and Molenberghs, 2014). We estimate a logit model of the probability of consent, using the mode of interview and controls for differences in sample composition as identified in Table 6: sex, academic qualifications, age, total personal income, employment status, and sample origin – whether they were part of the original IP sample, the IP4 refreshment sample, or the IP7 refreshment sample. Mode of interview remains a significant predictor of consent. In fact, controlling for differences in sample composition does not explain the difference in consent rates between face-to-face and web respondents: the average marginal effect of the mode coefficient in the logistic model is 20.6 percentage points, similar to the unconditional differences in consent rate by mode of interview. This method assumes that the selection into mode is fully explained by the covariates included. This assumption is unlikely to hold, so the estimate may not fully account for mode selection. As an aside, we find that there is no statistically significant difference between sample origin, so "time in panel" did not have an effect on consent.

The third method of controlling for selection effects accounts for differences in observed and unobserved characteristics between face-to-face and online respondents by using the random allocation to web-first versus CAPI-first as an instrument for the actual mode of interview. The design of the mixed modes experiment corresponds to an 'encouragement design with non-compliance to treatment' in clinical trials (Greenland, 2000): cases assigned to the treatment (web survey) can revert to the control treatment (face-to-face interview), however cases in the control treatment cannot revert to the treatment. The assumptions underlying this method require an instrumental variable that is correlated with the actual mode of interview, not correlated with the variables driving the selection into mode, and not correlated with consent. The random allocation to web or face-to-face first provides such an instrument. The mode coefficient estimated from an instrumental variable regression suggests that web respondents were 30.0 percentage points less likely to give consent than respondents interviewed face-to-face.

Table 7: Effect of mode on consent to link to FCA credit rating data

	CAPI		Web			
	% consent	N (base)	% consent	N (base)	Difference	P-value
Mode of interview	67.1	1,020	48.1	1,076	19.0	<0.001
Intention to treat (issued mode)	72.3	661	50.8	1,466	21.5	<0.001
Covariate adjustment	-	-	-	2,073	20.6	<0.001
Instrumental variable	-	-	-	2,096	30.0	<0.001

These results show that the estimated effect of mode on consent is in fact higher when selection is accounted for (30.0 percentage points with the instrumental variable regression) than when selection and measurement are confounded (19.0 percentage points by mode of interview). This reinforces the findings in Table 6, suggesting that respondents who complete the survey online are people who are generally more likely to consent. However, completing the survey online rather than with a face-to-face interviewer considerably reduces their probability of giving consent.

The pattern of results reported in Table 7 was replicated when we estimated mode differences in consent rates to link to DWP administrative data. Respondents who had not previously given consent to link DWP administrative data to their survey responses were also asked consent to DWP linkage in IP9. This included those who had once previously withheld their consent, and those who had turned 16 since the previous wave. Despite the differences in the eligible sample for the consent request, the results for the DWP linkage consent are similar to those we find for the FCA consent and therefore not reported.

3. Qualitative research: how do respondents decide whether to give consent to linkage?

The aim of the qualitative research was to investigate the mechanisms underpinning respondents' decisions to consent. Understanding how people make this decision helps understand the reasons why respondents' consent decisions can be inconsistent between topics and over time, and why the mode of interview has such a large effect on the consent decision.

3.1 Methods

We conducted in-depth face-to-face interviews with 25 members of the *Understanding Society* Innovation Panel. Interviews lasted 60 minutes and took place in March/April 2017 in participants' homes. Interviews were digitally recorded and participants received a £40 thank you voucher. To represent a range of views, participants were purposively selected from the Innovation Panel sample to include people who had given consent to data linkage, not given consent, or given consent one year but not the next (or vice versa), and to balance the mode in which they had completed their last interview (face-to-face or online). Selection was further based on location, gender, age group, educational qualifications, difficulty in meeting financial obligations each month, and whether or not participants received State benefits.

Interviewers used a semi-structured discussion guide and a range of stimulus materials and activities. Participants were shown two consent questions (in varying order), that have been used in the *Understanding Society* survey: consent to link economic data held by the DWP, and consent to link health data held by the NHS. Participants were asked for their responses and about their understanding of these questions. In addition, researchers used the projective technique of construction, to access thoughts or beliefs that are less conscious or more difficult to verbalise: participants were asked to assess the benefits and risks of data linkage from the perspective of a third person. See Beninger et al. (2017) for the discussion guide, the wording of the two consent questions and other materials used for the interviews.

To analyse the qualitative interview data we used an iterative and inductive approach, starting with the views expressed by participants, and included elements of 'grounded theory' analysis (Charmaz, 2006; Glaser and Strauss, 1967). We used a thematic framework (Ritchie et al., 2003) approach whereby findings from interviews are coded and organised thematically within a framework developed from the aims of the research. This method of synthesising enables drawing out both the diversity of opinions and common themes emerging from the interviews. The quotations from participants referenced in the text below are documented in Table 8.

3.2 Results

How do participants interpret consent questions?

Informed consent requires understanding four key elements of the linkage process: i) the direction of information flows, ii) the extent of information shared, iii) the parties involved, and iv) the relevance and purpose of data linkage. There was a lot of variation in how participants understood these elements, such that the initial, unprompted levels of understanding can be described as either accurate, confused, or inaccurate.

Participants whose understanding was accurate understood that the general purpose of data linkage was to obtain better data for research purposes. They also understood the direction of flow of their personal information (Q1, Q2). Only a few participants recognised that consent implied that past and future data about them could be linked (Q3).

Participants who were confused by the consent questions misunderstood key components of the request. Participants generally understood that data would be shared and that ISER would have access to their data held by the other organisation. They were however confused about whether people within those other organisations would have access to their survey data, whether they would use the data for purposes other than research, and whether third parties would have access to their linked data (Q4, Q5, Q6). A common view was that the actual survey data would be shared with government departments (Q7). Participants who were familiar with the data-holding organisation were more likely to comprehend what type of information that organisation has about them. Others however struggled to understand what information the administrative records contain (Q8). Participants were also unclear about the purpose of data linkage. Some thought ISER mistrusted the accuracy or truthfulness of the answers they had given in the survey. People who thought the organisation did not hold any data about them questioned the value of them consenting to the data linkage (Q9).

Participants whose understanding of the consent question was inaccurate tended to think that their survey responses would be shared with government departments for them to use (Q10). Participants who misinterpreted the consent request in this way, feared that data linkage was a way for government departments to monitor them.

Overall, understanding was improved if participants read the information booklet: those who read it were more likely to understand the direction and extent of data flows, and the purpose of data linkage. Understanding was however hampered by certain words and phrases (for details see Beninger, Digby and MacGregor 2017), and by the reading habits of some participants. Participants with limited literacy, dyslexia, or of older age said they tended to read small sections of the consent question that seemed important, usually just skimming the first and last sentences, rather than reading the full question text. They described this as habitual behaviour that they had adopted to manage the process of reading.

What do participants think are the implications of giving consent to linkage?

Overall, few participants understood what happens after they give consent: participants were vaguely aware that information is uploaded to a computer and used for research purposes.

Participants who understood the consent request did not identify any ways in which they would personally benefit from consenting (Q11). They did however identify third parties that would benefit. They generally understood that consent would lead to more accurate information about them and it was thought that this would benefit academics, because the information would be used for research purposes (Q12). Participants also thought that more accurate information would improve government policy planning, which might benefit society more generally in future (Q13). Some participants who misunderstood the flow of information however thought that consent could lead to improvement of services for them personally, for example in the provision of health care or pensions. Other participants who did not understand the purpose of data linkage saw no potential benefits to anyone (Q14).

Most participants recognised that consenting to data linkage entailed personal risks for them, that could either result from inadequate data protection in the processes of storing, handling and sharing information, or from misuse of personal data by government departments. The safety of personal information and the risks of identity fraud were prevalent concerns. While participants understood that websites and servers were encrypted, they thought that data are ‘never truly safe’, and can also be misplaced by accident (Q15, Q16). Stigma, embarrassment or reputational damage resulting from leaked data revealing sensitive or compromising personal information was also a concern. Some participants believed that sharing their contact details put them at risk of harassment by telemarketers (Q17). The risk that government departments could misuse personal information to 'catch out' individuals for participants for illegal or inappropriate behaviour was also a concern (Q18, Q19). Participants who believed in the possibility of personal risks were less likely to consent (Q20).

What influences the participant's decision whether or not to give consent?

Whether or not participants gave consent was not only determined by whether they had understood what the consent questions were asking. Instead the consent decision was based on a combination of subconscious, rational, social and environmental factors.

Subconscious factors included personality traits and related habitual behaviours. Participants who self-identified as private, suspicious persons focused on the risks of sharing their personal data (Q21). In contrast participants who self-identified as more open, trusting persons were dismissive of these risks (Q22). Some participants mentioned habitual behaviours, such as answering 'yes' or 'no' by default in response to consent-type questions (Q23).

Rational factors influencing the consent decision were dominated by the participants' assessment of the risks and benefits of consenting. Participants who consented were more likely to say that they 'have nothing to hide', for example health issues or financial circumstances that would make them feel vulnerable if their administrative data were accessed (Q24). Some were willing to consent because they believed that the survey organisation and the government already had so much information about them, that there was 'nothing they don't already have' (Q25). Similarly, some participants thought that having already agreed to participate in the survey, 'why stop here': their trust in the survey led them to agree to allow researchers to gather the most effective data possible (Q26).

Participants who consented also thought that the benefits of consent outweighed concerns about personal privacy or data protection (Q27). Generally, individuals who work in professions where they are aware of data exchange and data protection were more able to rationally assess the risk of consent. Some dismissed the risks, while others had heightened concerns that undermined their willingness to consent (Q28). Participants who were less likely to consent were worried about the potential personal repercussions of consenting, such as hacking and fraud (Q29, Q30). They were particularly concerned about NHS data if they had health problems that they thought could prevent them from getting jobs in the future (Q31). A few participants were also concerned about potential repercussions of data security breaches for their children or other people close to them (Q32).

Literacy levels and cognitive ability also influenced rational decision making: a few participants were not able to understand the question (as judged by the interviewer or communicated by the participant themselves) and therefore said they were unable to give consent.

Social factors influencing the consent decision included friends and family, and perceived norms around data sharing. Some participants said they would consult their spouse before giving consent. Others seriously considered reasons for not consenting, because they thought certain friends or family members would be hesitant (Q33). Participants were also influenced by different perceptions of established norms around data sharing, including the view that 'everyone shares everything already', and that 'in this day and age, you can never be too cautious' about sharing personal information because of the risks of identity theft, fraud, etc.

Environmental factors included trust in the organisation running this particular survey, and more broadly, stories of data security breaches reported in the media. The nature of interactions with the survey organisation, and how long participants had participated in the survey, helped establish trust: the organisation was seen as a reputable organisation which had 'too much riding on the survey to mishandle information'. Trust reduced participants' concerns about the risks of data linkage and supported the decision to consent. In contrast news stories in the media heightened participants' awareness of phishing scams, fraud, identity theft and data leaks, which in turn increased concern about consenting to data linkage.

How does the survey mode influence the decision to consent?

We asked participants who had participated online as well as those who had not whether and how responding online would influence their decision to consent. Participants gave several reasons why they would be less willing to give consent in an online survey. The difficulty of reading a lot of content on a computer screen was one prevalent reason: participants said they were less likely to fully read information presented on screen than they would in a face-to-face setting. This tendency to skim read online text would lead them to overlook key aspects of the consent request (Q34).

Concern about privacy and security of personal information shared online was another key reason why participants would be less likely to consent online. In contrast participants felt higher levels of trust in the security of their data in the face-to-face setting where the participant could see where their data were going as the interviewer recorded their responses (Q35, Q36). Only a minority of participants realised that the mode in which the consent question is asked does not alter the linkage procedures and would therefore not affect their willingness to consent (Q37).

The possibility of asking the interviewer clarification questions and obtaining reassurance was another reason why participants thought they would be more likely to give consent in the face-to-face setting. Although we know from previous research examining the behaviours of interviewers and respondents that respondents only rarely ask any questions about consent requests (Burton, Sala and Knies 2014), the qualitative interviews indicated that the idea of support was a motivating factor (Q38, Q39). The physical presence of the interviewer also increased the willingness to consent through perceived social pressure to conform (Q40).

Why do participants change their consent decision over time?

Participants who have given inconsistent responses to consent questions at different points in time were asked about the reasons why they had changed their mind. These were participants who had for example not consented in past panel interviews, but did say they would give consent in the qualitative interview. The reactions of these participants highlighted the role of subconscious, environmental and social factors in driving decisions to consent. Participants were generally unable to recall whether they had consented in the past. When the interviewer reminded them of their previous responses, some were surprised about their previous decision and unable to rationalise why they had changed their mind. A few participants commented that the change was probably arbitrary and dependent on how they were feeling on the day.

For participants who, on looking back at their past consent response, were able to post-rationalise why they might have changed their consent decision, the reasons were wide-ranging. Some indicated that they might have been more or less concerned about infringements on their privacy in the past. Some thought perhaps their understanding of the purpose of linking their personal data had changed. Others mentioned life events and changes in circumstances that had changed their attitudes towards societal benefits of sharing data (Q41).

Several participants said they were more cautious about giving consent due to the perception that the increased use of technology led to a rise in hacking (Q42, Q43). Others were surprised that they had consented in the past. These participants (who misunderstood the direction of data flows) claimed they would not have consented if they had known that their information would be shared with government departments that could affect their personal finances, such as the DWP and HMRC.

Finally, a further reason expressed for changing consent was trust in the survey, built up over years of taking part, and recognition that sharing information with ISER had not led to infringements on their privacy. Participants claimed that this was reassuring and made them more likely to consent (Q44).

4. Discussion

The empirical results presented in this chapter show that many respondents are inconsistent in the consent decisions they make – between topics and over time. Revoking consent, although rare, is not related to whether respondents will agree to later consent requests for other domains. Similarly, a large proportion of non-consenters do give consent – for the same or other domains – when asked again in later interviews. This inconsistency in decisions over time suggests that consent is not generally based on strongly held views about data linkage. This interpretation of the empirical results is supported by findings from the qualitative interviews. These illustrate that many participants do not or only partially understand the linkage request and the implications of consenting. The qualitative interviews also highlight that rational consideration of the implications, risks and benefits of consenting to linkage only influence the decision in part. In addition, subconscious, social and environmental factors play an important role in determining the consent decision.

The empirical results also show that the mode of data collection has a large effect on the willingness to consent to data linkage: respondents are much less likely to give consent when completing a survey online than in a face-to-face interview. Our analyses suggest this is clearly a mode effect rather than a selection effect; in fact, respondents who complete the survey online are people who are generally more likely to consent. The qualitative interviews offer some insights into why participants are so much more reluctant to consent in a web survey. Participants seem more concerned about data security when answering survey questions online than in a CAPI interview. Only a few seem to realise that the linkage and data handling procedures are exactly the same, regardless of the mode in which they give consent. They also seem less likely to read the (often lengthy) consent questions and associated information materials and therefore more likely to overlook key pieces of information. The presence of the interviewer also supports willingness to consent, both by providing opportunity to ask clarification questions and obtain reassurance, and through social pressure to conform.

The quantitative and qualitative findings have a number of practical implications and point to needs for additional research.

In terms of practical implications, the finding that decisions about consent do not appear to be strongly held suggests it is worth re-asking consent of those who declined at some earlier wave.

Doing so does not appear to be coercive; that is, the practice does not seem to persuade people to agree to something they don't want to do. A related implication is that re-asking consent of those who already consented may increase the number of non-consenters. Instead, providing an opportunity to revoke consent (as is done in *Understanding Society*) seems to give participants sufficient control over their decisions.

Another practical implication is that there is a clear need to explain the linkage process better to survey respondents (see also Thornby et al., 2017). This may be more of an issue in the web mode, where respondents are less likely to read materials and interviewers cannot compensate, by explaining the process to respondents or addressing questions or concerns they may have.

A related implication is that mode matters in the consent to administrative data linkages. This finding is unlikely to lead to a return to face-to-face surveys to maximize consent rates, but considering interviewer follow-up on non-consenters is one possibility to consider. The mode finding is an important consideration for those studies considering switching from an interviewer-administered mode to a self-administered one for some or all of the panel members. In multi-mode surveys it would be best to reserve consent requests for interviewer-administered instruments.

Our findings also identify a number of research needs related to administrative linkage consent. First, given the differential consent rates by mode, understanding the reasons for this difference and finding ways to overcome the lower consent rates on the web are a priority. One explanation is that the request is subject to social desirability response bias – respondents are agreeing to the consent request more in face-to-face to please the interviewer. An alternative explanation is that the interviewer is taking an active role in clarifying the process, addressing respondent concerns, and persuading the respondent to consent. It may also be that the mere presence of an interviewer conveys the legitimacy of the request in a way that is currently not done online. Understanding the role of the interviewer in the consent process, and studying the interviewer-respondent interaction during the consent process may identify effective strategies used by interviewers that could be duplicated online. Simply providing additional information to online respondents may not be sufficient if they do not read the material provided or click on links to access such information.

As an aside, there may be parallels here to the standardized versus conversational interviewing debate. As Schober and Conrad (1997) and others have found, conversational interviewing has advantages when complex questions are asked. This approach provides interviewers with an opportunity to convey uncertainty and for interviewers to clarify meaning and address concerns. The online version of consent is much like standardized interviewing, whereas (successful) interviewers may be much better at anticipating and addressing respondent concerns. Getting respondents to recognize when they need definitions for complex terms (or clarification of the consent request) online is a challenge.

As already noted, finding ways to encourage respondents to raise questions and get answers before making a consent decision may help increase how informed they are (or feel they are). Addressing the *informed* part of the informed consent process, especially in online survey administration, is a key issue for both research and practice.

Finally, we identified a number of research gaps in our review of prior literature. Our quantitative and qualitative findings suggest that these gaps may become even more important as the number and variety of consent requests increases in longitudinal studies, and as such studies increasingly make use of both interviewer- and self-administered modes of data collection. Understanding how and why people make decisions when multiple consent requests are made, and figuring out the best ways to ask consents for multiple domains is an important next step. Similarly, the observation from the qualitative research that respondents seemed more worried about the security of their data when asked online for their consent – given that the linkage process itself is identical – is intriguing, and again suggests that the context of the request matters. There is much more we need to know about the *process* of administrative data linkage consent in different modes, in addition to the *outcome* of that process.

Table 8: Selected quotations from qualitative interviews

Q1	*“They want access to your health records to help them with the answers that they [have and that] you [are] going to provided them on your health and wellbeing."* (Female, has consented, face-to-face)
Q2	*"From what I understand, they want to get information from the DWP to cross-reference that against the answers I've given you in the survey…"* (Female, has consented, web)
Q3	*“I get it. Permission is being given for access [to] historic info. This consent would hold going forward, until I wrote a letter saying I wanted to revoke [my] consent.”* (Male, has not consented, face-to-face)
Q4	*"It's odd, because I would have thought it would be the NHS giving you info and not the other way round...It’s not clear which individuals within the NHS would have it…if it was my GP I wouldn't mind because I have a good relationship (with him), but if it was others...not sure. I'd be happy to give my consent, but I would want to know...which individuals within the NHS would get it...administrative records...it's a bit vague what that means."* (Female, has consented, web)
Q5	*“Who is actually getting the information, what is it being used for, is it being kept, is it going to be with my name?”* (Female, changed consent, web)
Q6	*"The first thing you always want to know is who would use it; the other stuff is incidental almost."* (Male, has consented, face-to-face)
Q7	*"‘Administration’ sounds a bit name and address, whereas this [referring to the NHS consent question] means anything that is… [in my] file would be handed over."* (Female, has consented, web)
Q8	*"People will say what is an employment programme, do you mean work or voluntary work."* (Male, changed consent, web)
Q9	*"I'm not sure they would get an answer; I would love it see but the last time I went to them they (NHS) didn't have anything (his records)."* (Male, has consented, face-to-face)
Q10	*"They want to share all this info with DWP about my records and when I've been working and not working I suppose."* (Male, has consented, web)
Q11	*"Giving consent doesn't help me, it doesn't help us personally, you're doing it because you're happy to potentially provide a better result, and potentially a more accurate result.”* (Male, changed consent, web)
Q12	*“It would provide a wider picture. if I was the one doing the research I would want as much information as I could have.”* (Female, changed consent, web)
Q13	*"It gives the government something to think about, if you link it all together you can see, instead looking at statistics and try to number crunch this should work, you are actually getting a feel for how this is actually effecting people and how people are actually coping with it from real life people and not just a computer programme that tells you how people should be coping with it".* (Female, has consented)
Q14	*“I can’t see any benefits. How does it help planning in the community or whatever?”* (Female, changed consent, face-to-face)
Q15	*"The more information you share, the more chance of identity theft, whether you delete the information or not it will still be in the systems at some point.”* (Male, has consented, web)
Q16	*"You have to bear in mind each time you use data you have more chance of it being adversely used, it's a small risk but not an negligible one either because the impact can be*

	massive." (Male, changed consent, web)
Q17	*"I think in terms of information getting leaked it's more annoying, I think it's annoying when people call you up because they've got your phone number and they start calling you up about a dishwasher or something or PPI [Payment Protection Insurance]...I would be more concerned about my contact details than the other details."* (Female, changed consent, face-to-face)
Q18	*"Well, I don't know if he claimed benefits when he perhaps shouldn't have so there's always that risk of being found out…Is it an elaborate scheme to find out who's been swindling the system?"* (Female, changed consent, web)
Q19	*"I think personally if I said go ahead and then all of the sudden the dole office flicked through and my name came up, and then they would say oh will just check on all of her, what is she getting and where is she and I'll be the unlucky bugger that will happen to."* (Female, changed consent, face-to-face)
Q20	*"Even though research is for the greater good...people aren't interested in that, they're interested in themselves...as soon as you think about risk it's just easier to close it down."* (Female, changed consent, face-to-face)
Q21	*"My default is not to... I'd definitely, outside of the Understanding Society, say no but I trust them."* (Male, changed consent, web)
Q22	*"It isn't in my nature; I don't tend to [worry about those kinds of things]."* (Male, has consented, web)
Q23	*"I think I was just so in the habit, and I still am, of if a site asks you for information you don't give it, unless you trust the source."* (Female, changed consent, web)
Q24	*"I would say yes because there is nothing I would say that could be used against me. If I was doing something dodgy, working and claiming and avoiding tax and stuff like that I would say 'no, I don't want to do it'."* (Male, has consented, face-to-face)
Q25	*"I've already sold myself to the devil..."* (Female, changed consent, face-to-face)
Q26	*"If you're not prepared to give consent don't do Understanding Society it the first place, because it is a lot of personal information already."* (Male, changed consent, web)
Q27	*"I do regard it as for the greater good of society, if people do research society. I think it is a valid form of research … I'm happy to support it."* (Male, has consented, face-to-face)
Q28	*"I work in an environment where data protection and security is a big deal. My default position in my industry is why has that information been requested... if it doubt we are told to say no."* (Male, changed consent, web)
Q29	*"Not sure because it is to with the money I'm getting and there's a lot on the telly about them wanting to cut you down. Is this one of these things when they've got ya."* (Female, changed consent, face-to-face)
Q30	*"I don't know enough about Understanding Society, all the questions before have been multiple choice and not information as deep as that. Before and because there is so much of it (data theft) going around the world and even if Understanding Society were using it properly and in the right ways there's still so much hacking."* (Male, has consented, web)
Q31	*"Even when it comes to getting a job, there are certain things that you don't want to share, say if it was me applying for a government job in the future, then like, I'm not saying this is the case but say I had an alcohol problem in the past and then it came up in my records."* (Female, changed consent, web)
Q32	*"There is something that you want to keep private and if your records are open for a certain amount of time it links to the children. I know it is only for survey purposes and your*

	confidentiality is assured but how many times in the press recently had you seen people losing records. What are they going to do with it, probably nothing but I'm a private person anyway." (Male, changed consent, web)
Q33	*"Mates are worried about me doing the survey so they would likely question this request."* (Female, has not consented, face-to-face)
Q34	*"Because the impact isn't as great is it, if you read something online. It's like, you know, it's like reading an email at work - you might not take that information in. But when you're doing it face to face, and you're sitting there, you know, I'm going to read this, aren't I, because you just asked me to."* (Female, has not consented, face-to-face)
Q35	*"I guess I am more likely to consent in person because I don't know who is asking [questions] online."* (Male, changed consent, face-to-face)
Q36	*"They may think online it could go anywhere because we all know once it's online it is for everyone to see. I think face to face they would feel more comfortable."* (Female, changed consent, web)
Q37	*"It would be the same thing, it doesn't matter; like I said if it was laid out like this and I had a good understanding and it was bold like that...'; 'if it was face to face or on the internet I'd just be like go ahead."* (Female, changed consent, web)
Q38	*"Doing it online is much easier to say no because usually it is just click a box whereas if someone is talking to you about it they can either be persuasive depending on whether they are supposed to be doing that or alleviate the concerns that just can't be answered by reading on the screen. You never feel you are quite getting enough information to make an informed decision when you are doing it through a screen".* (Male, changed consent, web)
Q39	*"If (regular interviewer) had done it face to face she usually gives extra detail and sometimes she can tell by the blank look on my face. She does tend to go into more detail sometimes."* (Female, consented, face-to-face)
Q40	*"Subtle pressures you get in a face to face…you don't want to let that person down*." (Female, has consented, web)
Q41	*"[Decision to consent] would depend on what was happening at the time."* (Female, changed consent, face-to-face)
Q42	*"As time is moving on you see more breaches of personal security."* (Male, changed consent, web)
Q43	*"Now with a lot more people hacking companies through personal details and everything developed through technology you need to me more careful of what info you are passing on to people."* (Male, changed consent, web)
Q44	*"Perhaps I just sort of got to a point and realised you know I've been doing it for however many, 6 years, and everything's been fine thus far and they are using it for research purposes."* (Female, changed consent, web)

References

Al Baghal, T. (2016), “Obtaining Data Linkage Consent for Children: Factors Influencing Outcomes and Potential Biases.” *International Journal of Social Research Methodology*, 19 (6): 623-643.

Bates, N.A. (2005), “Development and Testing of Informed Consent Questions to Link Survey Data with Administrative Records.” In *Proceedings of the American Statistical Association*, May 12–15, 2005, Miami Beach, FL, pp. 3786–3793. Available at: http://ww2.amstat.org/sections/srms/Proceedings/ (accessed December 23, 2017).

Beninger, K., Digby, A., and MacGregor, J. (2017), "Understanding Society: How People Decide Whether to Give Consent to Link their Administrative and Survey Data". *Understanding Society Working Paper 2017-13*. Colchester: University of Essex.

Burton, J., Sala, E., and Knies, G. (2014), "Exploring the Role of Interviewers in Collecting Survey Respondents' Consent to Link Survey Data to Administrative Records." Presented at the 4th Panel Survey Methods Workshop Ann Arbor, MI.

Calderwood, L. (2016), “Asking for Consent in a Mixed Mode Study – Issues and Outcomes.” Presented at the Mixing Modes and Measurement Methods in Longitudinal Studies Workshop. London, CLOSER.

Calderwood, L. and Lessof, C. (2009), “Enhancing Longitudinal Surveys by Linking to Administrative Data.” In P. Lynn (ed.), *Methodology of Longitudinal Surveys*. Chichester: Wiley, pp. 55-72.

Charmaz, K. (2006), *Constructing Grounded Theory: a Practical Guide though Qualitative Analysis*. London: Sage.

Dahlhamer, J.M. and Cox, C.S. (2007), “Respondent Consent to Link Survey Data with Administrative Records: Results from a Split-Ballot Field Test with the 2007 National Health Interview Survey.” In *Proceedings of the Federal Committee on Statistical Methodology Research Conference*, November 5–7, 2007, Arlington, VA. Available at: https://fcsm.sites.usa.gov/reports/research/2007research/ (accessed December 23, 2017).

Das, M. and Couper, M.P. (2014), “Optimizing Opt-Out Consent for Record Linkage.” *Journal of Official Statistics*, 30 (3): 479-497.

Eisnecker, P.S. and Kroh, M. (2017), “The Informed Consent to Record Linkage in Panel Studies: Optimal Starting Wave, Consent Refusals, and Subsequent Panel Attrition.” *Public Opinion Quarterly*, 81 (1): 131-143.

Freedman, V.A., McGonagle, K., and Andreski, P. (2014), “The Panel Study of Income Dynamics Linked Medicare Claims Data.” *PSID Technical Series Paper #14-01*, Ann Arbor, MI: Survey Research Center, University of Michigan.

Fulton, J.A. (2012), *Respondent Consent to Use Administrative Data*. College Park, MD: University of Maryland, Unpublished PhD dissertation. Available at: https://drum.lib.umd.edu/handle/1903/13601 (accessed December 23, 2017).

Gates, G.W. (2011), "How Uncertainty about Privacy and Confidentiality is Hampering Efforts to More Effectively Use Administrative Records in Producing U.S. National Statistics." *Journal of Privacy and Confidentiality*, 3 (2): 3-40.

Glaser, B.G. and Strauss, A.L. (1967), *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago, IL: Aldine de Gruyter.

Greenland, S. (2000), "An Introduction to Instrumental Variables for Epidemiologists." *International Journal of Epidemiology,* 29: 722-729.

Groves, R.M. and Harris-Kojetin, B. (eds.) (2017), *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*. Washington, DC: The National Academies Press. doi: 10.17226/24652.

Jäckle, A., Gaia, A., Lessof, C., and Couper, M.C. (in press) "Improving survey measurement of household finances: A review of new technologies and data sources", in P. Lynn (ed.) *Advances in Longitudinal Survey Methodology*, Chichester: Wiley.

Jäckle, A., Lynn, P., and Burton, J. (2015), "Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response." *Survey Research Methods*, 9 (1): 55-70.

Knies, G. and Burton, J. (2014), "Analysis of Four Studies in a Comparative Framework Reveals: Health Linkage Consent Rates on British Cohort Studies Higher than on UK Household Panel Surveys." *BMC Medical Research Methodology*, 14: 125

Korbmacher, J.M. and Schroeder, M. (2013), "Consent when Linking Survey Data with Administrative Records: The Role of the Interviewer." *Survey Research Methods*, 7, 115–131.

Kreuter, F., Sakshaug, J.W., and Tourangeau, R. (2016), "The Framing of the Record Linkage Consent Question." *International Journal of Public Opinion Research,* 28(1): 142-152.

Lessof, C. (2009), "Ethical Issues in Longitudinal Surveys." In P. Lynn (ed.), *Methodology of Longitudinal Surveys*. Chichester: Wiley, pp. 35-54.

Mostafa, T. (2014), "Variation within Households in Consent to Link Survey Data to Administrative Records: Evidence from the UK Millennium Cohort Study." *International Journal of Social Research Methodology*, 19 (3): 355-375.

Mostafa, T. and Wiggins, R.D. (2018), "What Influences Respondents to Behave Consistently When Asked to Consent to Health Record Linkage on Repeat Occasions?" *International Journal of Social Research Methodology*, 21 (1): 119-134.

Pascale, J. (2011), "Requesting Consent to Link Survey Data to Administrative Records: Results from a Split-Ballot Experiment in the Survey of Health Insurance and Program Participation (SHIPP)." Study Series: Survey Methodology 2011-03. Washington, D.C.: U.S. Census Bureau.

Ritchie, J. and Lewis, J. (2003), *Qualitative Research Practice*: A Guide for Social Science Students and Researchers. London: Sage Publications.

Sakshaug, J.W., Couper, M.P., Ofstedal, M.B., and Weir, D.R. (2012), "Linking Survey and Administrative Records: Mechanisms of Consent." *Sociological Methods and Research,* 41(4): 535-569.

Sakshaug, J.W., Tutz, V., and Kreuter, F. (2013), "Placement, Wording, and Interviewers: Identifying Correlates of Consent to Link Survey and Administrative Data." *Survey Research Methods*, 7(2): 133-144.

Sala, E., Burton, J., and Knies, G. (2012), "Correlates of Obtaining Informed Consent to Data Linkage: Respondent, Interview and Interviewer Characteristics." *Sociological Methods and Research,* 41(3): 414-439.

Sala, E., Knies, G., and Burton, J. (2014), "Propensity to Consent to Data Linkage: Experimental Evidence on the Role of Three Survey Design Features in a UK Longitudinal Panel." *International Journal of Social Research Methodology,* 17(5): 455-473.

Schober, M.F. and Conrad, F.G. (1997), "Does Conversational Interviewing Reduce Survey Measurement Error?" *Public Opinion Quarterly* 61(4): 576-602.

Singer, E. (1978), “Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys.” *American Sociological Review*, 43: 144–162.

Singer, E. (2003), “Exploring the Meaning of Consent: Participation in Research and Beliefs about Risks and Benefits.” *Journal of Official Statistics*, 19, 273–86.

Thornby, M., Calderwood, L., Kotecha, M., Beninger, K., and Gaia, A. (2017), “Collecting Multiple Data Linkage Consents in a Mixed Mode Survey: Evidence and Lessons Learnt from Next Steps.” *Centre for Longitudinal Studies Working paper 2017/13*, London: Institute of Education.

University of Essex (2017a), “Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009.” [data collection]. 9th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-10

University of Essex (2017b), “Understanding Society: Innovation Panel, Waves 1-9, 2008-2016.” [data collection]. 8th Edition. UK Data Service. SN: 6849, http://doi.org/10.5255/UKDA-SN-6849-9

Vannieuwenhuyze, J.T.A., Loosveldt, G., and Molenberghs, G. (2014), "Evaluating Mode Effects in Mixed-Mode Survey Data using Covariate Adjustment Models." *Journal of Official Statistics*, 30(1): 1-21.

Weir, D.R., Faul, J.D., and Ofstedal, M.B., (2014), “The Power of Persistence: Repeated Consent Requests for Administrative Record Linkage and DNA in the Health and Retirement Study.” Presented at the 4th Panel Survey Methods Workshop. Ann Arbor, MI.

APPENDIX: Wording of *Understanding Society* data linkage consent questions

WAVE / DOMAIN	TEXT	INFORMATION LEAFLET?
1 – Education	We would also like to add information from your education records. Here is a permission form and information leaflet. Please read this, ask me any questions and sign the form if you are happy for us to do this.	Yes Leaflet 1
1 – NHS data and Central Register	Finally, we would like to add some information from administrative health records to the answers you have given. This leaflet gives you information about what we would like to do. Please read it, ask me any questions and sign the form if you are happy for us to do this.	Yes Leaflet 2
4 – State Benefits	Thank you for taking part in this study. Your answers are of great value and interest to many academic and social policy researchers who will make good use of them. We would like to add records held by the Department for Work and Pensions [if region = England/Scotland/Wales] / Social Security Agency [if region = Northern Ireland] , containing information they hold on your benefit claims and time on employment programs, to the responses you have given to this study. Linking the information will help us to get a fuller picture about who claims benefits or seeks help from the DWP [if region = England/Scotland/Wales] / SSA [if region = Northern Ireland]. All information will be used for research purposes only. It will be used by academic or policy researchers under restricted access arrangements which make sure that the information is used responsibly and safely. Please read this leaflet explaining how we would like to attach your DWP [if region = England/Scotland/Wales] / SSA [if region = Northern Ireland] records to the answers you have given in this study and let me know if you have any questions. Do you give permission for us to pass your name, address, sex and date of birth to the DWP [if region = England/Scotland/Wales] / SSA [if region = Northern Ireland] for this purpose?	Yes Leaflet 3
4 – Education (NPD) (Consent to link education data not given/missing or is a new entrant and born after 1983)	Information about your education is held by The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4]. This information includes details such as the schools or colleges you've attended, the courses you've studied and the qualifications you've gained as well as other characteristics such as attendance, suspensions or exclusions, whether you have/had a special educational need, or have ever been eligible for free school meals. We would like to add this information to your survey answers. This will create a more accurate picture of your life and experiences, and will be used for research and statistical purposes only. It will be kept confidential and names and addresses are never included in the results. No individual can be identified in reports. We would like to link	Yes Leaflet 1

	in information about your education to your survey responses on a yearly basis. To do this we need your permission to pass the minimum amount of your personal details (your full name, gender, date of birth and full address) to The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] so they can identify your individual education record. They will then create a link between this record and your survey responses before making your records anonymous. The education record will be placed on a secure website so that researchers can use it for research and statistical purposes. Further information about what will happen to your data can be found in this leaflet and can also be found on the *Understanding Society* website. You may withdraw your permission for The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] to disclose your information at any time. Details on how to do this can be found in the leaflet and on the website.	
4 – Education (NPD) (Consent to link education data given and born after 1983)	In the past, you provided permission for us to obtain your education information from The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4]. We would very much like to be able to continue with this data linkage on a yearly basis to create a more accurate picture of your life and experiences. Further information about what will happen to your data can be found in this leaflet and can also be found on the *Understanding Society* website. You may withdraw your permission for The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] to disclose your education information at any time. Details on how to do this can be found in the leaflet and on the website. On this basis, can we continue to obtain your education record from The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2] / The Department for Education and Skills [if SCHOOLCNTRY = 3] / The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] for the purpose of adding the data to your survey responses?	Yes Leaflet 1
4 – NHS data and Central Register	We would like to add some information from administrative health records to the answers you have given. This leaflet gives you information about what we would like to do. Please read it, ask me any questions and sign the form if you are happy for us to do this.	Yes Leaflet 2

5 – Education (NPD)	We are interested in linking information from various public records to your survey answers. Before we finish this year's interview, we just need to ask if you are happy for us to do so. Information about your education is held by The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4]. This information includes details such as the schools or colleges you've attended, the courses you've studied and the qualifications you've gained as well as other characteristics such as attendance, suspensions or exclusions, whether you have/had a special educational need, or have ever been eligible for free school meals. We would like to add this information to your survey answers. This will create a more accurate picture of your life and experiences, and will be used for research and statistical purposes only. It will be kept confidential and names and addresses are never included in the results. No individual can be identified in reports. We would like to link in information about your education to your survey responses on a yearly basis. To do this we need your permission to pass the minimum amount of your personal details (your full name, gender, date of birth and full address) to The Department for Education [if SCHOOLCNTRY = 1]/ the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] so they can identify your individual education record. They will then create a link between this record and your survey responses before making your records anonymous. The education record will be placed on a secure website so that researchers can use it for research and statistical purposes. Further information about what will happen to your data can be found in this leaflet and can also be found on the *Understanding Society* website. You may withdraw your permission for The Department for Education [if SCHOOLCNTRY = 1]/ the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] to disclose your information at any time. Details on how to do this can be found in the leaflet and on the website.	Yes Leaflet 1
5 – Education (HESA)	We would like to add records held by the Higher Education Statistics Agency, or HESA, to the answers you have given us. These records contain information on your higher education experiences, including the subjects you studied and the qualifications you received. All information will be used for research purposes only by academic or policy researchers under restricted access arrangements which make sure the information is used responsibly and safely.	No

	Do you give permission for us to pass your name, postcode, sex, date of birth and higher education institution to HESA for this purpose?	
5 – Employment/ self-employment	We would like to add records held by HM Revenue and Customs, or HMRC, containing information on your employment and self-employment history, your income, National Insurance contributions and tax credits. All information will be used for research purposes only by academic or policy researchers under restricted access arrangements which make sure that the information is used responsibly and safely. Please read this leaflet explaining how we would like to attach your HMRC records to the answers you have given in this study and let me know if you have any questions. Do you give permission for us to pass your name, address, sex and date of birth to the HMRC for this purpose?	Yes Leaflet 4
5 – NHS data and Central Register	We would like to add some information from administrative health records to the answers you have given. This leaflet gives you information about what we would like to do. Please read it, ask me any questions and sign the form if you are happy for us to do this.	Yes Leaflet 2
5 – Vehicles	We would like to collect more information about the types of motor vehicles that people use. Could you give me the registration number of the vehicle(s) so we can look up some details about your vehicle(s) already held by the DVLA?	No
6 – Education (NPD)	We are interested in linking information from various public records to your survey answers. Before we finish this year's interview, we just need to ask if you are happy for us to do so. Information about your education is held by The Department for Education [if SCHOOLCNTRY = 1] / the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4]. This information includes details such as the schools or colleges you've attended, the courses you've studied and the qualifications you've gained as well as other characteristics such as attendance, suspensions or exclusions, whether you have/had a special educational need, or have ever been eligible for free school meals. We would like to add this information to your survey answers. This will create a more accurate picture of your life and experiences, and will be used for research and statistical purposes only. It will be kept confidential and names and addresses are never included in the results. No individual can be identified in reports. We would like to link in information about your education to your survey responses on a yearly basis. To do this we need your permission to pass the minimum amount of your personal details (your full name,	Yes Leaflet 1

	gender, date of birth and full address) to The Department for Education [if SCHOOLCNTRY = 1]/ the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] so they can identify your individual education record. They will then create a link between this record and your survey responses before making your records anonymous. The education record will be placed on a secure website so that researchers can use it for research and statistical purposes. Further information about what will happen to your data can be found in this leaflet and can also be found on the *Understanding Society* website. You may withdraw your permission for The Department for Education [if SCHOOLCNTRY = 1]/ the Education Analytical Services [if SCHOOLCNTRY = 2]/ The Department for Education and Skills [if SCHOOLCNTRY = 3]/ The Department of Education and the Education and Skills Authority [if SCHOOLCNTRY = 4] to disclose your information at any time. Details on how to do this can be found in the leaflet and on the website.	
6 – NHS data and Central Register	We would like to add some information from administrative health records to the answers you have given. This leaflet gives you information about what we would like to do. Please read it, ask me any questions and sign the form if you are happy for us to do this.	Yes Leaflet 2
7 – Education (NPD/ILR)	We are interested in linking information from various public records to your survey answers. Before we finish this year's interview, we just need to ask if you are happy for us to do so. We would like to add information from your education record to your survey answers. This will create a more accurate picture of your life and experiences, and will be used for research and statistical purposes only. Your information is confidential and names and addresses will never be included in the results. We would like to link in information about your education to your survey responses on a yearly basis. If you decide that you no longer wish to take part in the study we will retain your education information up until that point. After this point we will make no further links. You can withdraw your consent at any time. Please read this leaflet explaining what we would like to do and ask me any questions. INTERVIEWER - PLEASE HAND LEAFLET INFORMATION ON ADDING EDUCATION RECORDS TO RESPONDENT AND ALLOW RESPONDENT TIME TO READ IT AND ASK QUESTIONS. In order to locate your education records we need to send the data holder some personal information. Do you give permission to pass your name, address, sex, date of birth and last school attended to The Department for Education and the Department for Business, Innovation and Skills, as applicable or a contractor conducting the linkage? [if SCHOOLCNTRY = 1] / Do you give permission to pass your postcode, date of birth, gender and last	Yes Leaflet 5

	school attended to the Scottish Government Learning Directorate or a contractor conducting the linkage? [if SCHOOLCNTRY = 2] / Do you give permission to pass your name, address, sex, date of birth and last school attended to The Department for Education and Skills or a contractor conducting the linkage? [if SCHOOLCNTRY = 3] / Do you give permission to pass your name, address, sex, date of birth and last school attended to The Department of Education and the Education and Skills Authority or a contractor conducting the linkage? [if SCHOOLCNTRY = 4] Once your education records have been added to your survey responses, we will remove all your personal information so that your answers are anonymous.	
7 – State Benefits	Thank you for taking part in this study. Your answers are of great value and interest to many academic and social policy researchers who will make good use of them. We would like to add records held by the Department for Work and Pensions [if REGION = England/Scotland/Wales] / Social Security Agency [if REGION = Northern Ireland] , containing information they hold on your benefit claims and time on employment programs, to the responses you have given to this study. Linking the information will help us to get a fuller picture about who claims benefits or seeks help from the DWP [if REGION = England/Scotland/Wales] / SSA [if REGION = Northern Ireland]. All information will be used for research purposes only. It will be used by academic or policy researchers under restricted access arrangements which make sure that the information is used responsibly and safely. Please read this leaflet explaining how we would like to attach your DWP [if region = England/Scotland/Wales] / SSA [if region = Northern Ireland] records to the answers you have given in this study and let me know if you have any questions. Do you give permission for us to pass your name, address, sex and date of birth to the DWP [if region = England/Scotland/Wales] / SSA [if region = Northern Ireland] for this purpose?	Yes Leaflet 3
8 – Employment/ Self-employment	We would like to add records held by HM Revenue and Customs, or HMRC, containing information on your employment and self-employment history, your income, National Insurance contributions and tax credits. All information will be used for research purposes only by academic or policy researchers under restricted access arrangements which make sure that the information is used responsibly and safely. Please read this leaflet explaining how we would like to attach your HMRC records to the answers you have given in this study and let me know if you have any questions. Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	Yes Leaflet 4

IP9 – FCA	You can help in another way. People's experiences of financial services would be useful to share with the regulator of financial services, the Financial Conduct Authority (FCA), which has an objective of protecting consumers. The FCA would like to link your answers from the survey, anonymously, to other data it can obtain from regulated firms as part of its ongoing remit to help protect consumers. In addition, selected aggregated date from your credit files will be included in the anonymised data set that is provided to researchers for research purposes only. But we would only link these two sets of data with your consent. Consent to sharing your data will not affect your involvement with any financial services providers or your ongoing participation in the study. The information will be treated in strict confidence and used for research and to help the regulator of financial services to protect consumers. Do you give your consent to sharing your data? HELP = We'd like to link your survey and credit history data in three ways. Importantly, this process will protect your personal data. First, we would pass the information which identifies you, but not your answers from the survey, to a credit reference agency (CRA). Second, we would pass the survey results to the FCA, but the information which identifies you will be removed. The CRA will then provide an anonymous extract from your credit file to the FCA who will match this data using anonymised, numerical identifiers with survey data which includes your borrowing and repayments as well as income estimates. Third, selected aggregated data from your credit files will be included in the anonymised data set that is provided to researchers for research purposes only. To repeat, agreeing to this linkage or declining to take part will have no impact upon your ability to borrow or any other interaction with financial services providers. If you have further questions regarding this you are welcome to contact: xxx, Chief of Market Research. Email: xxx	No