National Centre for Research Methods Working Paper

1/20

Do respondents using smartphones produce lower quality data? Evidence from the UK Understanding Society mixed-device survey

Dr Olga Maslovskaya, Prof Peter WF Smith, Prof Gabriele Durrant

NCRM
National Centre for Research Methods

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

Do respondents using smartphones produce lower quality data? Evidence from the UK Understanding Society mixed-device survey

Authors: Dr Olga Maslovskaya[1] (om206@soton.ac.uk), Prof Peter WF Smith[1], Prof Gabriele Durrant[1]

[1] *University of Southampton, Department of Social Statistics and Demography, School of Economic, Social and Political Sciences, Faculty of Social Sciences, Highfield Campus, Southampton, SO17 1BJ, UK.*

Abstract

We live in a digital age with high level of use of technologies. Surveys have started adopting technologies including smartphones for data collection. There is a move towards online data collection in the UK, including an ambition to collect 75% of household responses online in the UK 2021 Census. Major social household surveys in the UK have either transitioned to online data collection or are in the process of preparation for the transitioning. The Covid-19 pandemic forced rapid transitions to online data collection for many social surveys globally, with this mode of data collection being the only possibility at the moment. There are still concerns regarding allowing respondents to use smartphones to respond to surveys and not much is known about data quality produced by respondents using smartphones for survey completion in the UK context. This paper uses the first available in the UK, large scale mixed-device survey, Understanding Society Wave 8 where 40% of the sample were assigned to online mode of data collection. It allows comparison of data quality between different devices within the online mode of data collection with a special focus on smartphones. This analysis is very timely and fills the gap in knowledge.

Descriptive analysis and then various regressions are used depending on the outcome variables to study data quality indicators associated with different devices in the online part of the survey. The following data quality indicators are assessed: break-off rates, item nonresponse, response style indicators, completion times, differential reporting indicators including self-reporting of risky behaviours, and consent to data linkage. Comparisons to limited results available in the UK are drawn. The results suggest that even in the context of non-optimised for smartphone questionnaire, we should not be concerned about respondents using smartphones for future social surveys, even for longer surveys such as the Understanding Society, as break off rates are very low and data quality between devices is not very different.

Keywords: data quality, data quality indicators, mixed-device online surveys, online survey, cross-sectional context

Acknowledgements

This work was supported by the ESRC Secondary Data Analysis Initiative grant “Understanding survey response behaviour in a digital age: Mixed-device online surveys and mobile device use”, grant number ES/P010172/1 and by the ESRC National Centre for Research Methods, Research Work Package 1, grant number ES/L008351/1.

Data Statement

This study uses the following dataset: the Understanding Society main survey Wave 8. The Understanding Society data were obtained from the UK Data Archive (http://www.data-archive.ac.uk/). Data Citation for the Understanding Society data: University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2018). *Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. *11th Edition.* UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-12.

Introduction

For number of years globally social surveys are experiencing lowering response rates. At the same time survey costs of social surveys are increasing as data collection agencies are trying to address issues associated with nonresponse. Another trend which is observed in the UK and also globally is increased internet access and internet use as well as increased ownership of mobile devices such as smartphones and tablets. According to ONS (2018), 90% of adults in the UK used the internet daily in 2018 and according to Ofcom (2018), 87% of the UK households had an internet connection in 2018. Also, according to Ofcom (2017), 76% of adults owned a smartphone in the UK in 2017 in comparison to only 27% in 2011. All these numbers continue to rise. To respond to all these trends, data collection organisations globally are undergoing paradigm shift and moving towards online data collection. Mixed-mode designs were introduced in some major social surveys in the UK as a cost saving initiative. Some surveys such as the Understanding Society have already moved to mixed-mode design (Carpenter & Burton, 2017), other major social surveys such as the UK Labour Force Survey (LFS) are currently undergoing an experimental stage in preparation to transition to online mode of data collection (Finlay et al., 2018a; Finlay et al., 2018b). Once everything is ready for online data collection and infrastructure is in place, it is cheaper to collect data online. The Covid-19 pandemic forced rapid transitions to online data collection for many social surveys globally, with this mode of data collection being the only possibility. However, the question remains regarding data quality: would data quality suffer from this shift in data collection method? There are still concerns regarding allowing respondents to use smartphones when completing surveys. Tourangeau et al. (2018) highlighted three main areas which could be a source for concerns regarding data quality produced by respondents using smartphones: 1. screens on smartphones are smaller than those on tablets or PCs which could lead to larger response order effects or more superficial processing of the questions; 2. touch screen interface which could lead to item nonresponse or inadvertent choice of the wrong answer; 3. smartphone respondents may complete surveys in settings where other people or other distractions are present which could lead to less candid reporting of sensitive information or lower quality reporting overall. Also, lack of optimisation of questionnaires for smartphones caused concerns in researchers regarding quality of data produced by respondents using smartphones for survey completion. Lack of optimisation was reported to negatively influence data quality (see Mavletova & Couper, 2015; Young et al., 2014; Peterson et al., 2013; Lorch & Mitchell, 2014; Arn et al., 2015; Horwitz, 2016; Revilla et al., 2016 and others). Social surveys have gone through various stages in how they were treating smartphones. Historically, due to concerns regarding lower data quality produced by respondents using smartphones, smartphones were blocked when respondents attempted using them for survey completion (Maslovskaya et al., 2019). Examples of the surveys which had this approach to smartphones in the UK context are Understanding Society Innovation Panel Wave 7, and the European Social Survey Mixed-mode experiment (ESSMM). Later, the use of smartphones was discouraged, for example, in the Second Longitudinal Study for Young People in England (LSYPE2), the National Child Development Study (NCDS) 2013-2014, the Community Life Survey (CLS) 2014-2015, and Understanding Society Innovation Panel Waves 8 and 9, and this is still the case in some of the surveys such as the Understanding Society Wave 8. However, some of the surveys have adopted "mobile-first" design where the questionnaire is designed with a small screen in mind and, therefore, there are no issues associated with non-optimised versions of questionnaire any longer. Currently the Office for National Statistics (ONS) UK is preparing for transition of the LFS online and they adopted "mobile-first" design to the Labour Market Survey (LMS) which is used as a test for transformation of the LFS to mixed-mode design. It is very important to compare data quality between smartphones and other devices in order to establish whether these concerns regarding lower data quality obtained through smartphones have any grounds in the UK context. The Understanding Society Wave 8 provides an opportunity to conduct this timely analysis as it has a large online component with a sizable group of people who chose to respond using smartphones and also the questionnaires were not optimised for smaller screen sizes. The main research question for this analysis is: *Do respondents who use smartphones for survey completion produce lower quality data in comparison to those using desktops/laptops or tablets?*

The next section of the paper reviews existing literature. The third section discusses the Understanding Society Wave 8 survey which was used for the analysis followed by the Methodology section. The results section summarises findings obtained form the analysis of various data quality indicators and compares them with the limited existing findings in the UK context. This section also discusses potential endogeneity issue and how this issue was addressed. The final section discusses implications of the results for survey practice.

Background and Data Quality Indicators

Data quality indicators include measures related to coverage, measurement effects and nonresponse. To assess the quality of questionnaires, measurement errors and the quality of substantive answers by device used by respondents various data quality indicators can be used. Analysis of data quality indicators will allow comparison of the quality of data collected through different devices within one survey. The following nonresponse indicators are used by researches: break-off rates by device and item nonresponse by device. To assess risk for measurement error by device the following data quality indicators can be useful: length of interview, response style indicators such as straightlining, differential reporting, number of grid questions in questionnaire, speed per question, number of answering categories and other indicators. A number of studies was conducted in Germany, Netherlands, Russia, and the US context (Andreadis, 2015; Gummer & Rossmann, 2014; Stapleton, 2013; Mavletova, 2013; Toepoel & Lugtig, 2014; Lugtig & Toepoel, 2016; Guidry, 2012; Barlas & Thomas, 2015; McClain & Crawford, 2013; Baker-Prewitt & Miller, 2013; Revilla & Couper, 2017; Tourangeau et al., 2018; Antoun et al., 2017; McClain et al., 2012; Buskirk & Andrus, 2014; Schlosser & Mays, 2017; Mavletova & Couper, 2014; Mavletova & Couper, 2015; Struminskaya et al., 2015). The majority of these studies was observational in nature. However, some of them had an experimental design (for example, Tourangeau et al., 2018) or conducted meta-analysis (for example, Mavletova & Couper, 2015). The main concerns regarding lower quality of data produced by smartphones are: (i) that it would take longer to complete surveys for those using smartphone (Andreadis, 2015; Gummer and Russman, 2014 – their results were true for both optimised and non-optimised surveys, Buskirk & Andrus, 2014; de Bruijne & Wijnant, 2013; Keusch & Yan, 2017; Struminskaya et al., 2015), (ii) the likelihood of break-offs is higher for smartphones especially if questionnaires are not optimised (Mavletova & Couper, 2015 – meta-analysis), (iii) also item non-response is higher for smartphone users (Mavletova & Couper, 2014 and 2015; Struminskaya et al., 2015; Lugtig & Toepoel, 2016; Keusch & Yan, 2017), (iv) as well as tendency of response style behaviours such as primacy effects (Stapleton, 2013) and straightlining or non-differentiation (Guidry, 2012; Barlas & Thomas, 2015; McClain & Crawford, 2013; Bake-Prewitt & Miller, 2013). However, it is important to mention that other studies found no difference in straightlining by device used by respondents (Revilla & Couper, 2017;

Tourangeau et al., 2018; Antoun et al., 2017) or in primacy effects (Mavletova, 2013; Toepoel & Lugtig, 2014; Wells et al., 2013; Matthews et al., 2018) as well as in item nonresponse (McCain et al., 2012; Wells et al., 2013; Buskirk & Andrus, 2014; Toepoel & Lugtig, 2014; Schlosser & Mays, 2017; Tourangeau et al., 2018 - experimental design study). Overall, in all these studies the responses produced by smartphones appeared to be very similar to responses obtained by other devices and the differences which are found are not large (Couper et al., 2017; Tourangeau et al., 2018). However, it is important to mention that all these results came from a non-UK context and more work is needed to obtain more conclusive results regarding data quality produced by respondents using smartphones for survey completion generally but also specifically in the UK context.

Due to lack of suitable for the analysis data, data quality by device used by respondents in the UK context is still underresearched. Matthew et al. (2018) investigated the following five areas of data quality: missingness, satisficing, survey length, response accuracy, and social desirability bias. They assessed these five areas through the following data quality indicators in the context of young people of the age of 16-17 in England (they used LSYPE2 for the analysis): break-off rates, item non-response, consent to data linkage, straightlining, primacy effects, acquiescence effects, completion time, response validation, and self-reported risky behaviours (Matthews et al., 2018). Their results were very reassuring and found no evidence of differences by device in item non-response, consent to data linkage, straightlining, primacy effects or acquiescence effects, completion time, response validation or self-reported risky behaviours. The only indicator they found differences by devices was break-off rates with slightly higher break-off rate among smartphone respondents (4% in comparison to 1% for PCs and laptops and 2% for tablets). These results are reassuring but the analysis was conducted using sample of the young people of age 16-17 and it is important to conduct similar assessment of data quality on the general population in the UK in order to be able to conclude whether smartphone users produce lower or the same quality data when compared to other device users.

Hanson et al. (2018) conducted preliminary assessment of data quality in the Community Life Survey (CLS) (2016-2017) which is a survey of the UK general population and in which 505 respondents used a smartphone (7% of the sample). They presented their results at the NatCen-ESS ERIC-City Methodology Seminar Series. They assessed only a limited number of data quality indicators (break-off rates and length of questionnaire) and found that break-off rates (or drop-out rates) are significantly higher for smartphones (13% in comparison to 7% observed for PCs and laptops and tablets). No evidence was found for differences in completion time by device. One of their main conclusions was that there was a need to analyse data quality across other UK surveys in order to obtain more conclusive results regarding the quality of data produced by respondents using smartphones for survey completion in the UK context.

Maslovskaya (2020) conducted an assessment of data quality in the Understanding Society Innovation Panel Wave 9. Respondents were discouraged from using smartphones for survey completion. The total sample size for online mode of data collection was not large with the group which choose to use smartphone for survey completion having only 83 respondents for whom device was known (7.4% of the sample). Therefore, these results should be interpreted with caution and more work is required to obtain more conclusive results. There was no difference by devices found for completion time and for agreeableness. Differences by device was found for item nonresponse. However, the item nonresponse was either lower for smartphones or no difference to other devices was found depending on the variable that was assessed. Mixed results were found for differential reporting. However, the final results suggest that even where device effects were found in differential reporting assessment, the differences were due to selection effects rather than due to device effects. Differences were found for straightlining with respondents using smartphones having a higher tendency for straightlining. Differences in break-off rates were also found to be significant (7.2% break-off rate for smartphones in comparison to 1.6% for desktops/laptops and 2.0% for tablets). However, due to the very low frequencies of break-offs, more work needs to be done as these results are obtained in bivariate analysis context. Also, differences by device were found for consent to data linkage with respondents using smartphone for survey completion having higher probability of not giving consent to data linkage. For more details of this analysis see Maslovskaya (2020).

It is important to assess differences in data quality produced by different devices in a large scale mixed-device survey in the UK context.

Data

This paper employs the Wave 8 of the UK Understanding Society survey - the Household Longitudinal Study in the United Kingdom. The survey covers topics of health, work, education, income, family and social life to help understand the long term effects of social and economic change, as well as policy interventions. Therefore, it is very important to establish whether the quality of data is equally high irrespective of devices respondents chose to use for survey completion in online mode of the survey. In Wave 8, Understanding Society moved to mixed-mode design as a cost saving initiative and introduced online mode of data collection. Specifically, Wave 8 used a push-to-web mixed-mode design in which 40% of participants were initially invited to complete the questionnaire online. A further 40% were initially invited to complete a face-to-face interview (CAPI) but then given the opportunity to complete survey online if they had not completed it in CAPI mode. The remaining 20% were a ring-fenced sample and were only approached for a face-to-face interview and a random sample of households were part of this CAPI-only sample. For the remaining 80%, the allocation to initial mode was not random. A model was used to identify households which were more likely to take part online and those households were included into the 40% which were initially invited to complete the questionnaire online. More details about the mixed-mode design and the sample allocations can be found in Understanding Society (2020).

This specific wave has the advantage over previous waves by collecting 40% of households online and having a large sample of respondents who chose to use smartphones when responding to the survey.

Respondents were able to complete online and to choose the device they wanted to use for survey completion. However, the advanced letter said: "The survey is available online at the website shown below, so you can complete it at a time that's best for you, although it might be easier for you if you use a computer rather than a mobile device." (https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/mainstage/fieldwork-documents/wave-8/advance-communications/wave-8-letters.pdf). The respondents were still discouraged from using smartphones and the reason for this was that the questionnaire was not optimised for the small screens of mobile devices including smartphones. However, despite this discouragement, 902 (11%) respondents in the online sample still chose to use smartphone for survey completion.

Data collection for each wave is scheduled across a 24-months period, with data collection taking place annually. Wave 8 data collection took place between January 2016 and December 2017.

The *final analysis sample*, including only cases who responded online and for whom the device used for survey completion was known, contains 7,972 respondents.

Methodology

The main variable of interest is *device used by respondents for survey completion* variable, it contains three categories: PCs and laptops – 5,055 (63.4%), tablets – 2,015 (25.3%) and smartphones – 902 (11.3%). Respondents were not assigned to a device and therefore were free to choose any device they wished for survey completion. Therefore, endogeneity might be a problem in the analysis and will be addressed below.

Outcome measures

We examine seven types of data quality indicators: completion time, response style indicators (extreme responses, primacy effects, and straightlining), break-off rates, item nonresponse in individual questions, differential reporting, consent to data linkage and self-reporting of sensitive questions.

The first outcome variable is completion time. The completion time is calculated in minutes using difference between end of interview and start of interview times. Completions times were positively skewed, therefore, logarithmic transformations of the time were obtained and used for modelling.

Various response style indicators were obtained for four blocks of attitudinal variables. The following blocks were used: 8 questions regarding personal good qualities (4-point scale), 12 questions on General Health Questionnaire (GHQ) aspects (4-point scale), 4 questions on satisfaction (7-point scale), and 7 question block on what is important for an individual (5-point scale). The following response style indicators were obtained for all four blocks of variables: primacy effects, extreme responses, and straightlining. These response style indicators are the tendencies of respondents to take cognitive "short cuts" when completing questionnaires (Krosnick, 1991; Roberts et al, 2019). We expect them to be more prominent in respondents using smartphones due to the need to scroll horizontally and vertically in the context of non-optimised questionnaire of the Understanding Society. The following two indicators were modelled: extreme responses and straightlining. If the questionnaire is not optimised, there is a high likelihood that extreme responses might not be visible to respondents and therefore the respondents would need to scroll horizontally which would increase burden on respondents. To obtain the total number of extreme responses, all responses with highest value within blocks of variables used for the analysis were added to obtain the total number of extreme responses. To measure straightlining we take the average deviation between the current answer compared to the answer for the preceding question (Loosveldt et al., 2018). The higher the score, the lower the straighlning tendency. This score is then converted into a binary indicator by giving respondents in the bottom decile a value of 1 (high straightlining tendence) and the remaining respondents a value of 0.

The third type of outcome indicators is break-off rate. The break-off rates are calculated on the basis of a binary variable which has value 1 for partially productive interviews and fully productive interviews are assigned a value of 0. According to Stephanie Auty (2019)[1] who is the Understanding Society User Support and Training Officer, "a partial interview is where someone has completed up to the household finances section, but not finished the interview. We define this as a useable partial. If someone does not get as far as the household finances section then they get an unproductive outcome, and their data is not included in the dataset."

[1] Personal communication via email on 08 August 2019.

The forth type of outcome indicators used for the analysis is item nonresponse. It is important to remember the context of the Understanding Society when deciding on the approach to item nonresponse investigation. The reason that we did not assess item nonresponse across entire questionnaire was due to the specific characteristic of the Understanding Society which has very low item nonresponse. This can be explained by loyalty of the respondents to the survey given that this specific wave is already wave 8 of the survey and, therefore, respondents have been with the survey for number of years and are loyal to the survey. Therefore, we selected six variables with the highest item nonresponse and used them separately for the analysis. Table 1 contains details of the variables which were assessed for item nonresponse. For each variable from Table 1 an additional variable was created in which value 1 was added to respondents who had a missing value for this specific variable and 0 otherwise.

Table 1: Variables used for item nonresponse assessment

Question in survey	Sample size	Missing values (percentages)
Should UK remain a member of the EU?	7,972	235 (2.9%)
Can I please have your home landline number?	1,845	63 (3.4%)
And can I please have your personal mobile phone number?	1,377	290 (21.1%)
Can I have a work phone number?	2,241	153 (6.8%)
Please enter your e-mail address here	1,801	26 (1.4%)
Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	3,280	30 (0.9%)

The fifth type of data quality indicators used for the analysis is differential reporting. These indicators help us to assess whether different devices might be associated with differences in reporting in different binary survey variables. Nine variables were selected for the analysis of differential reporting in the Understanding Society. Table 2 presents variables used for this type of data quality indicators assessment and associated sample sizes.

Table 2: Variables used for differential reporting assessment

Question in survey	Sample size
And are you male or female?	7,972
Are you in paid employment?	7,967
Do you have any long-standing physical or mental impairment, illness or disability? By ‘long-standing’ I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months.	7,960
Do you smoke cigarettes?	7,967
Do you regard yourself as belonging to any particular religion?	7,961
Should UK remain a member of the EU?	7,477

We would like to use your email address to keep in touch. What is your email address?	1,790
Do you normally have access to a car or van that you can use whenever you want to?	6,952
Do you have a full UK driving licence?	1,270

The sixth type of data quality indicators is a consent to data linkage. A variable regarding Her Majesty's Revenue and Customs (HMRC) record linkage (see Table 4.6) is used to assess differences in consent by device. This variable is a binary variable with value 1 attached when respondent does not want to give consent to data linkage and value 0 when respondent has read leaflet and is happy to give consent.

Finally, the seventh type of data quality indicators used is self-reporting of risky behaviours. Only a very small subset of the sample – youth self-completion questionnaire – was asked these sensitive questions (see Table 3). Binary variables were created with value 1 attached to any reports of risky behaviours irrespective of frequency where applicable and value 0 when risky behaviour was not reported.

Table 3: Variables used for assessment of self-reporting risky behaviours.

Question in survey	Sample size
Have you ever had an alcoholic drink? That is a whole drink, not just a sip	526
Thinking back over the last four weeks, how many times (if any) have you had five or more drinks on one occasion?	409
On how many occasions during the last 4 weeks (if any) have you been intoxicated or drunk from drinking alcohol, for example, staggered when walking, not being able to speak properly, throwing up or not remembering what happened?	407
In the last 12 months, have you tried cannabis (also known as marijuana, dope, hash or skunk)?	522
And any other illegal drug (including ecstasy, cocaine, speed)?	522
Since last interview, how many times you used or taken any illegal drugs?	523

Tables 4.1-4.7 present distributions of all outcome indicators used in the analysis by device used to complete the survey.

Table 4.1: Completion time indicator by device type

Completion time	PCs/Laptops	Tablets	Smartphones
	5007	1990	885
median (IQR)	35 (24)	37 (24)	36 (46)

Table 4.2: Response style indicators by device type

Response style indicators	PCs/Laptops	Tablets	Smartphones
Block 1	350	65	105
Primacy effect			
0	135 (38.6%)	32 (49.2%)	54 (51.4%)
1	77 (22.0%)	11 (16.9%)	11 (10.5%)
2	47 (13.4%0	2 (3.1%)	9 (8.6%)
3	36 (10.3%)	9 (13.8%)	13 (12.4%)
4	45 (12.9%)	8 (12.3%)	10 (9.5%)
5	6 (1.7%)	2 (3.1%)	3 (2.9%)
6	3 (0.9%)	1 (1.5%)	0 (0.0%)
8	1 (0.3%)	0 (0.0%)	5 (4.8%)
High extreme responses			
0	172 (49.1%)	31 (47.7%)	66 (62.9%)
1	41 (11.7%)	9 (13.8%)	11 (10.5%)
2	46 (13.1%)	5 (7.7%)	13 (12.4%)
3	43 (12.3%)	6 (9.2%)	4 (3.8%)
4	48 (13.7%)	11 (16.9%)	10 (9.5%)
5	0 (0.0%)	2 (3.1%)	0 (0.0%)
8	0 (0.0%)	1 (1.5%)	1 (1.0%)
Straightlining			
Straighlining tendency	72 (20.6%)	20 (30.8%)	35 (33.3%)
No straighlining tendency	278 (79.4%)	45 (69.2%)	70 (66.7%)
Block 2	5010	1996	885
Primacy effect			
0	1131 (22.6%)	454 (22.7%	231 (26.1%)
1	775 (15.5%)	327 (16.4%)	150 (16.9%)
2	572 (11.4%)	235 (11.8%)	116 (13.1%)
3	538 (10.7%)	224 (11.2%)	90 (10.2%)
4	511 (10.2%)	213 (10.7%)	79 (8.9%)
5	605 (12.1%)	207 (10.4%)	76 (8.6%)
6	645 (12.9%)	249 (12.5%)	89 (10.1%)
7	119 (2.4%)	38 (1.9%)	29 (3.2%)
8	43 (0.9%)	21 (1.1%)	12 (1.4%)
9	29 (0.6%)	8 (0.4%)	6 (0.7%)
10	26 (0.5%)	9 (0.5%)	4 (0.5%)
11	8 (0.2%)	5 (0.3%)	3 (0.3%)
12	9 (0.2%)	5 (0.3%)	3 (0.3%)
High extreme responses			
0	4620 (92.2%)	1833 (91.8%)	769 (86.9%)
1	176 (3.5%)	83 (4.2%)	54 (6.1%)
2	70 (1.4%)	19 (1.0%)	19 (2.1%)
3	39 (0.8%)	18 (0.9%)	11 (1.2%)
4	19 (0.4%)	11 (0.6%)	10 (1.1%)
5	18 (0.4%)	5 (0.3%)	1 (0.1%)
6	9 (0.2%)	6 (0.3%)	8 (0.9%)
7	21 (0.4%)	5 (0.3%)	0 (0.0%)
8	14 (0.3%)	4 (0.2%)	3 (0.3%)

9	8 (0.2%)	2 (0.1%)	2 (0.2%)
10	6 (0.1%)	2 (0.1%)	1 (0.1%)
11	5 (0.1%)	3 (0.2%)	4 (0.5%)
12	5 (0.1%)	5 (0.3%)	3 (0.3%)
Straightlining			
Straightlining tendecy	1196 (23.9%)	502 (25.2%)	211 (23.8%)
No straightlining tendency	3814 (76.1%)	1494 (74.8%)	674 (76.2%)
Block 3	5027	1997	890
Primacy effect			
0	4710 (93.7%)	1876 (93.9%)	826 (92.8%)
1	219 (4.4%)	74 (3.7%)	37 (4.2%)
2	57 (1.1%)	24 (1.2%)	15 (1.7%)
3	20 (0.4%)	11 (0.6%)	5 (0.6%)
4	21 (0.4%)	12 (0.6%)	7 (0.8%)
High extreme responses			
0	3534 (79.3%)	1410 (70.6%)	690 (77.5%)
1	778 (15.5%)	303 (15.2%)	114 (12.8%)
2	390 (7.8%)	141 (7.1%)	36 (4.0%)
3	189 (3.8%)	86 (4.3%)	21 (2.4%)
4	136 (2.7%)	57 (2.9%)	29 (3.3%)
Straightlining			
Straightlining tendency	1602 (31.9%)	672 (33.7%)	291 (32.7%)
No strraightlining tendency	3425 (68.1%)	1325 (66.3%)	599 (67.3%)
Block 4	5005	1984	884
Primacy effect			
0	987 (19.7%)	369 (18.6%)	155 (17.5%)
1	1425 (28.5%)	631 (31.8%)	300 (33.9%)
2	893 (17.8%)	332 (16.7%)	138 (15.6%)
3	802 (16.0%)	307 (15.5%)	121 (13.7%)
4	415 (8.3%)	156 (7.9%)	80 (9.0%)
5	270 (5.4%)	96 (4.8%)	52 (5.9%)
6	128 (2.6%)	44 (2.2%)	15 (1.7%)
7	85 (1.7%)	49 (2.5%)	23 (2.6%)
High extreme responses			
0	3767 (75.3%)	1424 (71.8%)	749 (84.7%)
1	739 *14.8%)	301 (15.2%)	73 (8.3%)
2	234 (4.7%)	114 (5.7%)	25 (2.8%)
3	96 (1.9%)	69 (3.5%)	7 (0.8%)
4	86 (1.7%)	34 (1.7%)	13 (1.5%)
5	30 (0.6%)	9 (0.5%)	4 (0.5%)
6	16 (0.3%)	13 (0.7%)	5 (0.6%)
7	37 (0.7%)	20 (1.0%)	8 (0.9%)
Straightlining			
Straightlining tendency	1200 (24.0%)	440 (22.2%)	204 (23.1%)

No straightlining tendency	3805 (76.0%)	1544 (77.8%)	680 (76.9%)

Table 4.3: Break-off rates by device type

Break off rates	PCs/Laptops	Tablets	Smartphones
	5055	2015	902
Partial interview	17 (0.3%)	10 (0.5%)	9 (1.0%)
Full interview	5038 (99.7%)	2005 (99.5%)	893 (99.0%)

Table 4.4: Item nonresponse by device type

Item nonresponse	PCs/Laptops	Tablets	Smartphones
Should UK remain a member of the EU?	5055	2015	902
Nonresponse	136 (2.7%)	74 (3.7%)	25 (2.8%)
Response	4919 (97.3%)	1941 (96.3%)	877 (97.2%)
Can I please have your home landline number?	1182	407	256
Nonresponse	44 (3.7%)	14 (3.4%)	5 (2.0%)
Response	1138 (96.3%)	393 (96.6%0	251 (98.0%)
And can I please have your personal mobile phone number?	887	338	152
Nonresponse	197 (22.2%)	75 (22.2%)	18 (11.8%)
Response	690 (77.8%)	263 (77.8%)	134 (88.2%)
Can I have a work phone number?	1374	522	345
Nonresponse	91 (6.6%)	42 (8.0%)	20 (5.8%)
Response	1283 (93.4%)	480 (92.0%)	325 (94.2%)
Please enter your e-mail address here	1190	437	174
Nonresponse	17 (1.4%)	8 (1.8%)	1 (0.6%)
Response	1173 (98.6%)	429 (98.2%)	173 (99.4%)
Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	2040	818	422
Nonresponse	16 (0.8%)	9 (1.1%)	5 (1.2%)
Response	2024 (99.2%)	809 (98.9%)	417 (98.8%)

Table 4.5: Differential reporting by device type

Differential reporting	PCs/Laptops	Tablets	Smartphones
And are you male or female?	2509	774	292
Male	2509 (49.6%)	774 (38.4%)	292 (32.4%)

Female	2546 (50.4%)	1241 (61.6%)	610 (67.6%)
Are you in paid employment?	5052	2013	902
Yes	3020 (59.8%)	1271 (63.1)	735 (81.5%)
No	2032 (40.2%)	742 (36.9%)	167 (18.5%)
Do you have any long-standing physical or mental impairment, illness or disability? By 'long-standing' I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months.	5045	2014	901
Yes	1654 (32.8%)	647 (32.1%)	217 (24.1%)
No	3391 (67.2%)	1367 (67.9%)	684 (75.9%)
Do you smoke cigarettes?	5052	2013	902
Yes	464 (9.2%)	213 (10.6%)	138 (15.3%)
No	4588 (90.8%)	1800 (89.4%)	764 (84.7%)
Do you regard yourself as belonging to any particular religion?	5050	2013	898
Yes	2451 (48.5%)	976 (48.5%)	319 (35.5%)
No	2599 (51.5%)	1037 (51.5%)	579 (64.5%)
Should UK remain a member of the EU?	4762	1865	850
Remain	2880 (60.5%)	1015 (54.4%)	530 (62.4%)
Leave	1882 (39.5%)	850 (45.6%)	320 (37.6%)
We would like to use your email address to keep in touch. What is your email address?	1184	434	172
Yes	1045 (88.3%)	360 (82.9%)	151 (87.8%)
No	139 (11.7%)	74 (17.1%)	21 (12.2%)
Do you normally have access to a car or van that you can use whenever you want to?	4407	1791	754
Yes	4118 (93.4%)	1698 (94.8%)	709 (94.0%)
No	289 (6.6%)	93 (5.2%)	45 (6.0%)
Do you have a full UK driving licence?	790	293	187
Yes	193 (24.4%)	78 (26.6%)	49 (26.2%)
No	597 (75.6%)	215 (73.4%)	138 (73.8%)

Table 4.6: Consent to data linkage by device type

Consent to data linkage	PCs/Laptops	Tablets	Smartphones
Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	2020	808	417
Happy to give consent	628 (31.1%)	225 (27.8%)	141 (33.8%)
Does not want to give consent	1392 (68.9%)	583 (72.2%)	276 (66.2%)

Table 4.7: Self-reporting of risky behaviour by device type

Self-reporting of risky behaviour	PCs/Laptops	Tablets	Smartphones
Have you ever had an alcoholic drink? That is a whole drink, not just a sip	345	66	106
Yes	306 (86.4%)	57 (86.4%)	95 (89.6%)
No	48 (13.6%)	9 (13.6%)	11 (10.4%)
Thinking back over the last four weeks, how many times (if any) have you had five or more drinks on one occasion?	270	51	88
None	81 (30.0%)	18 (35.3%)	18 (20.5%)
Once	65 (24.1%)	15 (29.4%)	30 (34.1%)
Twice	54 (20.0%)	8 (15.7%)	18 (20.5%)
3-5 times	50 (18.5%)	7 (13.7%)	17 (19.3%)
6-9 times	16 (5.9%)	3 (5.9%)	2 (2.3%)
10+ times	4 (1.5%)	0 (0.0%)	3 (3.4%)
On how many occasions during the last 4 weeks (if any) have you been intoxicated or drunk from drinking alcohol, for example, staggered when walking, not being able to speak properly, throwing up or not remembering what happened?	268	51	88
0	134 (50.0%)	30 (58.8%)	44 (50.0%)
1-2	94 (35.1%)	18 (35.3%)	29 (33.0%)

3-5	31 (11.6%)	2 (3.9%)	11 (12.5%)
6-9	6 (2.2%)	1 (2.0%)	3 (3.4%)
10 or more	3 (1.1%)	0 (0.0%)	1 (1.1%)
In the last 12 months, have you tried cannabis (also known as marijuana, dope, hash or skunk)?	352	66	104
Yes	64 (18.2%)	7 (10.6%)	17 (16.3%)
No	288 (81.8%)	59 (89.4%)	87 (83.7%)
And any other illegal drug (including ecstasy, cocaine, speed)?	352	66	104
Yes	27 (7.7%)	4 (6.1%)	8 (7.7%)
No	325 (92.3%)	62 (93.9%)	96 (92.3%)
Since last interview, how many times you used or taken any illegal drugs?	352	66	105
Never	285 (81.0%)	58 (87.9%)	83 (79.0%)
Once or twice	27 (7.7%)	5 (7.6%)	11 (10.5%)
3 or 4 times	9 (2.6%)	1 (1.5%)	4 (3.8%)
5-10 times	14 (4.0%)	1 (1.5%)	0 (0.0%)
More than 10 times	17 (4.8%)	1 (1.5%)	7 (6.7%)

Data analysis

Majority of data quality indicators used for the analysis are binary variables. Completion time and primacy effects and extreme responses indicators are continuous. A logarithmic transformation of the completion time variable is normally distributed. The types of indicators identified bivariate tests of associations and then appropriate regression modelling technique applied for the analysis.

For the bivariate analysis, Chi-squared tests of associations were used for binary and multinomial data quality indicators and Kruskal-Wallis test was used for completion time due to the variable's positive skewness as this test is used for differences in distributions of skewed continuous variables between groups.

Linear regression was used to model the logarithmic transformation of the completion time indicator as well as some of the response style behaviour indicators such as extreme responses. For all other indicators binary logistic regression was used for the analysis. As all respondents are clustered within households, robust standard errors were estimated to control for potential clustering (Huber, 1967; White, 1980, 1984 and 1994). STATA version 13 software package was used for the analysis.

Literature (for example, Hox et al., 1991 and Steinbrecher et al., 2015 (break-offs); Yan & Tourangeau, 2008 (completion time)) suggests that various data quality indicators are associated with respondents' demographic and socio-economic characteristics. The following demographic and socio-economic characteristics of respondents are used as independent variables in the analyses: gender, age, qualification, employment status, economic activity, income, marital status, household type, children in household, ethnicity, country of residence, government office region (GOR), urban/rural. Table 5 contains all demographic and socio-economic characteristic of respondents by device type employed in the analysis.

Table 5: Characteristics of respondents by device type

Characteristics	PCs or laptops	Tablets	Smartphones
Sex	5055	2015	902
Male	2509 (49.6%)	774 (38.4%)	292 (32.4%)
Female	2546 (50.4%)	1241 (61.6%)	610 (67.6%)
Age	5055	2015	902
16-19	250 (4.9%)	40 (2.0%)	65 (7.2%)
20-29	506 (10.0%)	171 (8.5%)	216 (23.9%)
30-39	631 (12.5%)	322 (16.0%)	253 (28.0%)
40-49	875 (17.3%)	365 (18.1%)	208 (23.1%)
50-59	1046 (20.7%)	433 (21.5%)	102 (11.3%)
60-69	1022 (20.2%)	426 (21.1%)	44 (4.9%)
70+	725 (14.3%)	258 (12.8%)	14 (1.6%)
Marital status	5038	2006	896
Single	1288 (25.6%)	410 (20.4%)	370 (41.3%)
Married or in a civil partnership	3073 (61.0%)	1293 (64.5%)	421 (47.0%)
Separated or divorced	484 (9.6%)	229 (11.4%)	97 (10.8%)
Widowed	193 (3.8%)	74 (3.7%)	8 (0.9%)
Ethnicity	5021	2003	894
White UK	4487 (89.4%)	1839 (91.8%)	797 (89.1%)
Irish and any other White	233 (4.6%)	55 (2.7%)	29 (3.2%)
Mixed	56 (1.1%)	12 (0.6%)	11 (1.2%)
Indian, Pakistani, Bangladeshi	133 (2.6%)	41 (2.0%)	41 (4.6%)
Chinese and other Asians	43 (0.9%)	21 (1.0%)	6 (0.7%)
African and Carribean	45 (0.9%)	26 (1.3%)	5 (0.6%)
Other ethnicities	24 (0.9%)	9 (0.4%)	5 (0.6%)
Highest qualification	5047	2013	901
Degree	1908 (37.8%)	557 (27.7%)	292 (32.4%)
Other higher degree	706 (14.0%)	317 (15.7%)	120 (13.3%)
A-levels	1058 (21.0%)	405 (20.1%)	257 (28.5%)
GCSEs	816 (16.2%)	443 (22.0%)	174 (19.3%)
Other qualification	332 (6.6%)	162 (8.0%)	38 (4.2%)
No qualification	227 (4.5%)	129 (6.4%)	20 (2.2%)

Responsible for children under 16	5055	2015	902
Yes	565 (11.2%)	302 (15.0%)	259 (28.7%)
No	4490 (88.8%)	1713 (85.0%)	643 (71.3%)
Current economic activity	5052	2015	902
Self-employed	489 (9.7%)	145 (7.2%)	69 (7.6%)
Paid employment	2489 (49.3%)	1097 (54.4%)	639 (70.8%)
Unemployed	122 (2.4%)	42 (2.1%)	33 (3.7%)
Retired	1391 (27.5%)	542 (26.9%)	32 (3.5%)
Maternity leave	18 (0.4%)	21 (1.0%)	17 (1.9%)
Family care and home	122 (2.4%)	63 (3.1%)	32 (3.5%)
Full time student	287 (5.7%)	39 (1.9%)	58 (6.4%)
Long-term sick or disabled	82 (1.6%)	44 (2.2%)	12 (1.3%)
Doing something else, including apprenticeship, unpaid work for family business etc.	52 (1.0%)	22 (1.1%)	10 (1.1%)
In paid employment	5052	2013	902
Yes	3020 (59.8%)	1271 (63.1%)	735 (81.5%)
No	2032 (40.2%)	742 (36.9%)	167 (18.5%)
Household type	5055	2015	902
1 adult, no children	694 (13.7%)	263 (13.1%)	90 (10.0%)
1 adult with children	77 (1.5%)	41 (2.0%)	41 (4.5%)
Couple or 2 adults, no children	2124 (42.0%)	879 (43.6%)	194 (21.5%)
Couple or 2 adults with children	997 (19.7%)	427 (21.2%)	331 (36.7%)
3 or more adults, no children	775 (15.3%)	268 (13.3%)	148 (16.4%)
3 or more adults with children	388 (7.7%)	137 (6.8%)	98 (10.9%)
Total gross monthly personal income	5052	2015	902
1st income quartile	1178 (23.3%)	440 (21.8%)	227 (25.2%)
2nd income quartile	1030 (20.4%)	465 (23.1%)	173 (19.2%)
3rd income quartile	1206 (23.9%)	528 (26.2%)	254 (28.2%)
4th income quartile	1638 (32.4%	582 (28.9%)	248 (27.5%)
Country of residence	5051	2015	902
England	4398 (87.1%)	1729 (85.8%)	781 (86.6%)
Wales	180 (3.6%)	77 (3.8%)	35 (3.9%)
Scotland	356 (7.0%)	141 (7.0%)	59 (6.5%)
Northern Ireland	117 (2.3%)	68 (3.4%)	27 (3.0%)
GOR	5051	2015	902
North East	194 (3.8%)	94 (4.7%)	54 (6.0%)
North West	481 (9.5%)	242 (12.0%)	110 (12.2%)
Yorkshire and the Humber	497 (9.8%)	196 (9.7%)	100 (11.1%)
East Midlands	399 (7.9%)	184 (9.1%)	69 (7.6%)

West Midlands	462 (9.1%)	186 (9.2%)	80 (8.9%)
East of England	561 (11.1%)	203 (10.1%)	97 (10.8%)
London	443 (8.8%)	155 (7.7%)	70 (7.8%)
South East	881 (17.4%)	300 (14.9%)	124 (13.7%)
South West	480 (9.5%)	169 (8.4%)	77 (8.5%)
Wales	180 (3.6%)	77 (3.8%)	35 (3.9%)
Scotland	356 (7.0%)	141 (7.0%)	59 (6.5%)
Northern Ireland	117 (2.3%)	68 (3.4%)	27 (3.0%)
Urban or rural area	5051	2015	902
Urban	3900 (77.2%)	1583 (78.6%)	767 (85.0%)
Rural	1151 (22.8%)	432 (21.4%)	135 (15.0%)

Endogeneity

This study is an observational study and does not have an experimental design. This is one of the limitations of the survey used for the analysis. There is a concern (as for any other observational study) that when we observe device effects that they are really selection effects. For example, are observed differences between responses to questions on employment after controlling for potential confounders due to the device or due to differences in unobserved characteristics across the device types, the so-called omitted variables problem? In an observational study device type maybe correlated with the error term and therefore it is endogenous. Different disciplines have various approaches to addressing the issues of endogeneity. For example, in econometrics instrumental variable approach is commonly used in the context of double-hurdle or Heckman selection models (Heckman, 1979). Instrumental variables are variables that are correlated with the endogenous variable but uncorrelated with the error term conditional on other covariates (Greene, 2012). We tried to identify a strong instrumental variable but it was a difficult in the specific context of the analysis. Another approach commonly used in epidemiology when experimental design is unavailable is Propensity Score Matching (Bai & Clark, 2019). It allows "imitation" of an experimental design and helps to address potential endogeneity problem. This approach estimates the probability or a propensity score of being in the "treatment" group. It then matches individuals from the "treatment" and "control" groups based on these scores. There are various methods for matching available. An appropriate method is then used to test for treatment effect in the matched sample. However, the main limitation of this approach for our context is that most methods of propensity score matching can only handle two groups and in our context we are interested in contrasting three groups (desktop/laptop, tablet and smartphone). Another limitation of this approach is that during matching procedure potentially a large proportion of sample might be lost due to lack of available matches which then makes it difficult to infer the results to the wider population of interest. Biologists use so called negative controls approach (Lipsitch et al., 2010). They identify "control" responses which are assumed not to be related to the "treatment". Appropriate tests are then used to test for treatment effects. "Spurious" treatment effects imply selection, otherwise we can conclude that there is indeed a treatment effect (or in our case a device effect). For example, we assume that there is no device effect on reporting gender. Therefore, if reported gender is associated with device used by respondents after controlling for other covariates, we then can conclude that we observe an omitted variable issue and can conclude that there is a selection effect and not a device effect. This approach is used in this analysis to address potential endogeneity problem.

Results

Bivariate analysis

Tests of bivariate association (Kruskal-Wallis test and Chi-squared test) were used to identify which outcome indicators are associated with devices used by respondents for survey completion in a bivariate context. Table 6 shows the results of bivariate associations by outcome indicators.

Table 6: Results of bivariate associations by outcome indicators

Outcome indicator		Bivariate association
Completion time		Yes ***
Response style indicators		
	Block 1	
	Primacy effect	No
	High extreme responses	Yes**
	Straightlining	Yes*
	Block 2	
	Primacy effect	Yes**
	High extreme responses	Yes***
	Straightlining	No
	Block 3	
	Primacy effect	No
	High extreme responses	Yes***
	Straightlining	No
	Block 4	
	Primacy effect	No
	High extreme responses	No
	Straightlining	No
Break-off rates		Yes *
Item Nonresponse		
	Should UK remain a member of the EU?	No
	Can I please have your home landline number?	No
	And can I please have your personal mobile phone number?	No
	Can I have a work phone number?	No
	Please enter your e-mail address here	No
	Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	No
Differential Reporting		

And are you male or female?	Yes***
Are you in paid employment?	Yes***
Do you have any long-standing physical or mental impairment, illness or disability? By 'long-standing' I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months.	Yes***
Do you smoke cigarettes?	Yes***
Do you regard yourself as belonging to any particular religion?	Yes***
Should UK remain a member of the EU?	Yes***
We would like to use your email address to keep in touch. What is your email address?	Yes*
Do you normally have access to a car or van that you can use whenever you want to?	No
Do you have a full UK driving licence?	No
Consent to data linkage	
Do you give permission for us to pass your name, address, sex and date of birth to HMRC for this purpose?	No
Self-reporting of Risky Behaviour	
Have you ever had an alcoholic drink? That is a whole drink, not just a sip	No
Thinking back over the last four weeks, how many times (if any) have you had five or more drinks on one occasion?	No
On how many occasions during the last 4 weeks (if any) have you been intoxicated or drunk from drinking alcohol, for example, staggered when walking, not being able to speak properly, throwing up or not remembering what happened?	No
In the last 12 months, have you tried cannabis (also known as marijuana, dope, hash or skunk)?	No
And any other illegal drug (including ecstasy, cocaine, speed)?	No
Since last interview, how many times you used or taken any illegal drugs?	No

Note: *-p<0.05; **-p<0.01; ***-p<0.001

The results from descriptive analysis suggest that there is no association between device used by respondents for survey completion and self-reported risky behaviours or consent to data linkage. These indicators will not be further assessed in the modelling context. These results are consistent with the findings by Matthews et al. (2018). However, Maslovskaya (2020) found an association in consent to data linkage and device used but different variable (consent to data linkage of benefits data) was used for the previous analysis in the context of Innovation Panel.

Majority of the variables in the survey have very low item nonresponse. The results of descriptive analysis of the variables with highest item nonresponse rates suggest that for all six variables there is no significant association between item nonresponse and device used by respondents. These results are also consistent with results by Matthews et al. (2018). However, Maslovskaya (2020) found association between item nonresponse and device in Innovation Panel of Understanding Society but the direction of association was the opposite to one expected, i.e. likelihood of item nonresponse was lower for smartphones than for other devices.

Kuruskal-Wallis test results suggest that there is a significant difference in completion times by devices with completion times being longer for tablets and smartphones when compared to PCs and laptops.

The results for response style indicators are mixed and also consistent with earlier finding by Matthews et al. (2018). There is a significant association between device used and primacy effect for block 2. There is a consistent association between device used and extreme responses across all blocks of variables used for the analysis with the exception for block 4. Also, there is an association between straightlining and device used found in block 1. All these significant associations will be further analysed in the modelling context.

Chi-squared test suggests that there is a significant association between break-off rates and devices used with smartphones having the highest rate of break-offs (1% in comparison to 0.3% for desktops and laptops and 0.5% for tablets). However, it is very important to mention here that only 36 respondents had partially productive interviews. Therefore, these very low frequencies do not allow us to investigate break-off rates by devices further in the modelling context. The break-off rates are also consistently very low, possibly due to the loyalty of the respondents to the survey. These results are consistent with results reported by Matthews et al. (2018), Hanson et al. (2018) and Maslovskaya (2020). However, the break-off rates were reported to be higher in other surveys for all devices with the highest group being smartphone the same as in the main survey of Understanding Society (LSYPE2 – 3.5%; CLS – 13%; Innovation Panel – 7.2%).

When differential reporting is assessed, the majority of variables used for the analysis is bivariately associated with the device used by respondents for survey completion with an exception for the questions regarding access to a car or a van and driving licence. The variables for which significant association with device was found will be further investigated in modelling context with the hope that once we have controlled for other variables, the device effect which is observed in the bivariate context will disappear. Modelling results are summarised and presented in Table 7.

Table 7: Modelling results

Outcome indicator	Significance in modelling context	Controlling for variables
Completion time	Yes	Device is still significant after controlling for age, race, household type, income, health and GOR. It takes 18% longer on average to complete the survey using smartphone than PC. No difference between desktop and tablet

Response style indicators		
Block 1		
High extreme responses	No	After controlling for sex
Straightlining	Yes	Device used is significant after controlling for sex. Straightlining is more likely in tablets and smartphones when compared to desktops and laptops
Block 2		
Primacy effect	No	After controlling for sex
High extreme responses	No	After controlling for sex and age
Block 3		
High extreme responses	No	After controlling for sex and age
Differential Reporting		
And are you male or female?	Yes	Device used is still significant after controlling for income, household type, economic activity, qualification, marriage, age. Higher likelihood to be a woman in tablets and smartphones than in PCs.
Are you in paid employment?	Yes	Device used is still significant after controlling for GOR, income, age, martial status, race, children, household type. Higher likelihood of being not employed in PCs than in smartphones or tablets
Do you have any long-standing physical or mental impairment, illness or disability? By 'long-standing' I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months.	No	After controlling for age, marital status, race, household type and economic activity
Do you smoke cigarettes?	No	After controlling for age, marital status, race, qualification, household type, income, GOR, urban, sex and economic activity
Do you regard yourself as belonging to any particular religion?	No	After controlling for age, marital status, race, income, GOR and sex
Should UK remain a member of the EU?	No	After controlling for age, race, qualification, household type, income, GOR, sex
We would like to use your email address to keep in touch. What is your email address?	Yes	Device used is still significant after controlling for marital

		status, race, qualification, household type, income and sex. Higher likelihood of not having email is in those with tablets in comparison to those with PCs but no difference between PCs and smartphones.

In the modelling context, completion time variable (ln(min)) was still found to be significantly associated with device used by respondents after controlling for various significant variables: age, race, household type, income, health and GOR. It takes 18% longer on average to complete the survey using smartphones than PCs or laptops. No difference in completion times was found between PCs or laptops and tablets. These results contradict the findings by Matthews et al. (2018), Hanson et al. (2018) and Maslovskaya (2020) where in other three contexts no evidence of difference in time taken to complete the survey by device was found.

This finding is not surprising and expected. There are two plausible explanations for the effect observed. The first one is that the questionnaire was not optimised and, therefore, the task for those using smartphones was harder due to potential need to scroll horizontally and vertically to see all possible answers to the questions and due to other possible difficulties with non-optimised questionnaire and, therefore higher burden on respondents. The second explanation might be due to the fact that it might have taken longer to complete each question for those using smartphone due to contextual effects as they might have been multi-tasking or being distracted. This hypothesis can be tested by comparing response times by individual questions between devices and will be done as a part of further work.

When different response style indicators were modelled, no associations were found to be significant between either extreme responses or primacy effects and device used by respondents. However, in block 1 the straightlining indicator was still found to be significantly associated with device used by respondents after controlling for respondent's gender variable. The results suggest that the straightlning is more likely in tablets and smartphones when compared to desktops or laptops. This finding was also not surprising and might be explained by the fact that the questionnaire was not optimised for mobile devices. For attitudinal questions, there is a higher likelihood of all answers not being visible on small screen and, therefore, would require horizontal scrolling which might cause straightlining in some of the respondents as a short-cut to quicker completion of a burdensome task. Maslovskaya (2020) reported similar results in Innovation Panel of the Understanding Society for association between straightlining tendency and device used by respondents.

As for differential reporting, after controlling for possible confounding factors, we still found an association between device used and three actual survey variables (gender, employment status and access to an email). Higher likelihood of reporting being a woman was observed in tablets and smartphones when compared to PCs and also higher likelihood of being not employed in PCs or laptops when compared to smartphones and tablets. Also, there was a higher likelihood of not having an email in those using tablets than PCs but there was no difference in likelihood of not having an email in those using smartphones than PCs or laptops. After applying negative controls approach to a potential issue of endogeneity, we believe that in all three cases we observe selection effect or omitted variable issue rather than a device effect. The same results were reported by Maslovskaya (2020) in the analysis of Innovation Panel for gender and employment variables. From previous analysis by Maslovskaya et al. (2019) we know that women are more likely to use smartphones as well as those being employed. In the third variable, we did not find difference between smartphones and PCs and as we know that tablet users are on average older than other device users, there is a higher likelihood for them not to have an email address. We are confident that the results presented above suggest selection effects rather than device effects and are reassuring as they contribute to the evidence that we can be confident about allowing respondents using smartphones for survey completion even when questionnaires are not optimised for mobile devices.

Discussion

Data collection organisations are currently undergoing a paradigm shift and more and more social surveys are conducted online either exclusively or as a part of mixed-mode designs. Some social surveys have already moved to mixed-mode design in the UK such as the Understanding Society survey. Other surveys are still at the stage of testing and preparation for transitions online such as the UK LFS. In the past respondents were only allowed to use desktops or laptops for survey completion. However, due to the increase in ownership of mobile devices such as tablets and smartphones, it is clear that blocking those devises would not benefit data collection and data quality. Therefore, more and more surveys either optimise their questionnaire for smartphones or adopt mobile-first designs for questionnaires. Due to differences in screen sizes between different devices, there are still concerns regarding allowing respondents to use mobile devices and specifically smartphones for survey completion as smartphones are suspected to produce lower quality data, especially in the non-optimised questionnaire context. A number of studies were conducted in different survey and country contexts to gather evidence regarding relative to other devices quality of data produced by smartphones. However, still not much is known about data quality obtained from smartphone survey completion in the UK context due to lack of suitable data for the analysis. Currently, the UK is preparing for the 2021 UK Census and it is planned to collect returns from 75% of households online.

Therefore, it is crucially important to address this gap in knowledge in the UK context. The Wave 8 of the Understanding Society provides a unique opportunity to address the issues of data quality produced by respondents using smartphones for survey completion as it has a large online sample with a large group of respondents who choose to use smartphone for survey completion. The following data quality indicators were assessed: completion time, response style indicators (straightlining, primacy effects, and extreme responses), item nonresponse, break-off rates, differential reporting, consent to data linkage, and self-reporting of risky behaviours.

The results of the analysis suggest that although the Understanding Society Wave 8 survey was not optimised for smartphones, the data collected from respondents using smartphones for survey completion seem to be of similar quality as data collected through other devices such as PCs and laptops. No difference by devices were found for primacy effects, for extreme response, for item nonresponse, consent to data linkage or self-reporting of risky behaviours. Break-off rates were found to be significantly different by devices with higher rates observed for smartphones. However, the break-off rate for smartphone was around 1.0% which is very low when compared to the results obtained by meta-analysis conducted by Mavletova & Couper (2015). They reported that on average break-off rates for smartphone users are around 14.0%. Therefore, the results obtained in this study are very reassuring for the smartphone group. Differential reporting analysis suggested that there were differences in reporting some of the actual survey variables by device used by respondents. However, by using negative controls approach to address potential endogeneity issue we concluded that the observed associations in gender and employment variables are there due to selection effects or omitted variables rather than device effects. The main differences between smartphone and other devices come from completion time with smartphone users taking longer to complete questionnaire and with higher straightlining tendency for respondents using smartphones for survey completion. We hope that questionnaire optimisation or mobile-first approach to questionnaire design will remove completion time and response style indicators differential effects. The following waves of the Understanding Society data should allow for testing this hypothesis as respondents will not be discouraged from using mobile devices any longer. All these results are very reassuring and suggest that survey practitioners and data collection organisations should be confident about allowing respondents to use smartphones for survey completion and for Census 2021 data collection as the data quality by devices in the UK context is not very different.

References

Arn, B., Klug, S. & Kolodziejski, J. (2015). Evaluation of an adapted design in a multi-device online panel: A DemoSCOPE case study. *Methods, Data, Analyses*, *9 (2)*, 185-212.

Andreadis, I. (2015). Web surveys optimized for smartphones: Are there differences between computer and smartphone users?. *Methods, Data, Analyses*, *9*(2), 213-228.

Antoun, C., Couper, M.P. & Conrad, F.G. (2017). Effects of mobile versus PC web on survey response quality: A Crossover experiment in a probability web panel. *Public Opinion Quarterly*, *81 (5)*, 280-306.

Bai H. & Clark M.H. (2019). *Propensity score methods and applications*. Sage Publications.

Baker-Prewitt, J. & Miller, J. (2013). *Mobile research risk: What happens to data quality when respondents use a mobile device for a survey designed for a PC*. Paper presented at the CASRO Online Research Conference, San Francisco, US.

Barlas, F.M. & Thomas, R.K. (2015). *On the go: How mobile participants affect survey results*. Paper presented at the European Survey Research Association (ESRA) conference, Reykjavik, Iceland.

Buskirk, T.D. & Andrus, C. (2014). Making mobile browser surveys smarter: Results from a randomized experiment comparing online surveys completed via computer or smartphone. *Field Methods*, *26(4),* 322-342.

de Bruijne, M. & Wijnant, A. (2013). Comparing survey results obtained via mobile devices and computers: An Experiment with a mobile web survey on a heterogeneous group of mobile device versus a computer-assisted web survey. *Social Science Computer Review, 31 (4),* 482-504.

Carpenter H. & Burton, J. (2017). Moving Understanding Society to mixed mode: Effects on response and attrition. Paper presented at the Understanding Society Conference, University of Essex, Colchester, UK.

Couper, M.P, Antoun, C. & Mavletova, A. (2017). Mobile web surveys. In P.P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E.Lyberg, N.C.Tucker and B.T. West (eds.), *Total survey error in practice* (pp.133-154). New Jersey: John Wiley & Sons.

Finlay, S., Stannard, J., Silvester, H. & Daxon, L. (2018a). *Labour Market Survey Response Rate Experiments: Report for Test 1: Materials experiment*. Ipsos MORI Social Research Institute, from https://gss.civilservice.gov.uk/wp-content/uploads/2018/04/Test-1_Full-report_FINAL-for-publishing.pdf

Finlay, S., Stannard, J. & Daxon, L. (2018b). *Labour Market Survey Response Rate Experiments: Report for Test 2, Tranche 1: Incentives experiment*. Ipsos MORI Social Research Institute, from https://gss.civilservice.gov.uk/wp-content/uploads/2018/04/Test-2-Tranche-1-report-FINAL-for-publishing.pdf

Greene, W.H. (2012). *Econometric analysis* (chapter 8), 7th Edition. Pearson.

Guidry, K.R. (2012). *Response Quality and Demographic Characteristics of Respondents Using a Mobile Device on a Web-based Survey*. Paper presented at the AAPOR conference, Orlando, FL. Available at http://cpr.indiana.edu/uploads/Response%20Quality%20and%20Demographic%20Characteristics%20of%20Respondents.pdf

Gummer, T. & Roßmann, J. (2015). Explaining interview duration in web surveys: A multilevel approach. *Social Science Computer Review*, *33(2),* 217-234.

Hanson, T., Matthews, P. & McGee, A. (2018). *Completing social surveys on smartphones: What should we be worried about?* Paper presented at the NatCen-ESS ERIC-City Methodology Seminar Series, London, UK.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153–61.

Horwitz, R. (2016). *The Differential Effect of a Mobile-Friendly Instrument on Data Quality*. Paper presented at the PAPOR conference, San Francisco, UK.

Hox, J. J., de Leeuw, E. D. & Kreft, G. G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In P.P.Biemer, R.M.Groves, L.E.Lyberg, N.A.Mathiowetz & S.Sudman (eds.), *Measurement errors in surveys* (pp. 439-461). New York: John Wiley & Sons.

Huber, P.J. (1967). The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA: University of California Press, 1*, 221–233. Available at https://projecteuclid.org/euclid.bsmsp/1200512988.

Keusch, F. & Yan, T. (2017). Web versus mobile web: An Experimental study of device effects and self-selection effects. *Social Science Computer Review 35 (6)*, 751-769.

Krosnick, J. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology*, *5*, 213–236.

Lipsitch M., Tchetgen T.E. & Cohen T. (2010) Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology, 21*, 383–88.

Loosveldt, G, Wuyts, C. & Beullens, K. (2018). Interviewer variance and its effects on estimates. *Quality Assurance in Education*, *26 (2)*, 227-242.

Lorch, J. & Mitchell, N. (2014). *Why you need to make your surveys mobile friendly now*. Paper presented at a webinar, Shelton, CT: Survey Sampling International, webinar slides. http://www.websm.org/uploadi/editor/doc/1470902373Lorch_Mitchell_2014_Why_you_need_to_make_your_surveys.pdf

Lugtig, P. & Toepoel, V. (2016). The use of PCs, smartphones, and tablets in a probability-based panel survey: Effects on survey measurement error. *Social Science Computer Review*, *34 (1)*, 78-94.

Maslovskaya, O. (2020) *Data quality in mixed-device online surveys in the UK*. Paper presented at the City, University of London/ European Social Survey HQ/ NatCen Social Research Survey Methodology Seminar Series. Webinar, 30 April 2020. Available at https://www.youtube.com/watch?v=pT2wygY_OeI&feature=youtu.be.

Maslovskaya, O., Durrant, G., Smith P.W.F., Hanson, T. & Vilar, A. (2019). What do we know about mixed-device online surveys: Evidence from six UK surveys. *International Statistical Review 87(2)*, 326-346.

Matthews, P., Bell, E. & Wenz, A. (2018). Surveying young people in the smartphone age. *Social Research Practice 5,* 2-11.

Mavletova, A. (2013). Data quality in PC and mobile web surveys. *Social Science Computer Review*, *31(6)*, 725-743.

Mavletova, A. & Couper, M.P. (2014). Mobile Web Survey Design: Scrolling versus Paging, SMS versus E-mail Invitations. *Journal of Survey Statistics and Methodology, 2 (4)*, 498-518.

Mavletova, A. & Couper, M P. (2015). A Meta-Analysis of Breakoff Rates in Mobile Web Surveys. In D. Toninelli, R. Pinter & P. de Pedraza (eds.), *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies* (pp.81-98). London: Ubiquity Press.

McClain, C.M., Crawford, S.D. & Dugan, J.P. (2012). *Use of mobile devices to access computer-optimized web surveys: implications for respondent behavior and data quality*. Paper presented at the AAPOR Conference. Orlando, FL.

McClain, C. & Crawford, S.D. (2013). *Grid formats, data quality, and mobile device use: A questionnaire design approach*. Paper presented at the AAPOR Conference, Boston, MA.

Ofcom (2017). *Communications market report: United Kingdom*. From https://www.ofcom.org.uk/__data/assets/pdf_file/0017/105074/cmr-2017-uk.pdf.

Ofcom (2018). *Communications market report: United Kingdom*. From https://www.ofcom.org.uk/__data/assets/pdf_file/0022/117256/CMR-2018-narrative-report.pdf.

ONS (2018). *Internet users, UK: 2018*. From file:///C:/Users/om206/Downloads/Internet%20users,%20UK%202018.pdf.

Peterson, G., Mechling, J., LaFrance, J., Swinehart, J. & Ham, G. (2013). *Solving the unintentional mobile challenge*. Paper presented at the CASRO Online Research Conference, San Francisco, US.

Revilla, M. & Couper, M.P. (2017). Comparing grids with vertical and horizontal item-by-item formats for PCs and smartphones. *Social Science Computer Review*, *36 (3)*, 349-368.

Revilla, M., Toninelli, D., Ochoa, C. & Loewe, G. (2016). Do online access panels really need to allow and adapt surveys to mobile devices? *Internet Research, 26 (5),* 1209-1227.

Roberts, C., Gilbert, E., Allum, N. & Eisner, L. (2019). *Satisficing in surveys: A systematic review of the literature*. Paper presented at the Conference of the European Survey Research Association (ESRA), Lausanne, Switzerland.

Schlosser, S. & Mays, A. (2017). Mobile and dirty: Does using mobile devices affect the data quality and the response process of online surveys? *Social Science Computer Review, 36 (2)*, 212-230.

Stapleton, C.E. (2013). The smartphone way to collect survey data. *Survey Practice*, 6 *(2),* 1-7.

Steinbrecher, M., Roßmann, J. & Blumenstiel, J. E. (2015). Why do respondents break off web surveys and does it matter? Results from four follow-up surveys. *International Journal of Public Opinion Research*, *27(2),* 289-302.

Struminskaya, B., Weyandt, K. & Bosnjak, M. (2015). The effects of questionnaire completion using mobile devices on data quality. Evidence from a probability-based general population panel. *Methods, Data, Analyses*, *9 (2)*, 261-292.

Toepoel, V. & Lugtig, P. (2014). What happens if you offer a mobile option to your web panel? Evidence from a probability-based panel of Internet users. *Social Science Computer Review, 32(4),* 544-560.

Tourangeau, R., Sun, H., Yan, T., Maitland, A., Rivero, G. & Williams, D. (2018). Web surveys by smartphones and tablets: Effects on data quality. *Social Science Computer Review*, *36 (5),* 542-556.

Understanding Society (2020). Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009, *User Guide*, April 2020, Colchester: University of Essex. From https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/mainstage/user-guides/mainstage-user-guide.pdf

Wells, T., Bailey, J. & Link, M. (2013). Filling the void: Gaining a better understanding of tablet-based surveys. *Survey Practice, 6 (1)*, 1-10.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*, 817–830.

White, H. (1984). *Asymptotic theory for econometricians*. Orlando, FL: Academic Press.

White, H. (1994). *Estimation, inference and specification analysis*. New York: Cambridge University Press.

Young, R., Crawford, S.D., Couper, M.P. & Nelson, T. (2014). *The Effect of Mobile Web Survey Design on Screen Orientation Manipulation*. Paper presented at the MAPOR, Chicago, US.

Yan, T. & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22(1),* 51-68.