Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: coverage and participation rates and biases

Annette Jäckle (University of Essex)
Jonathan Burton (University of Essex)
Mick P. Couper (University of Michigan)
Carli Lessof (University of Southampton)

Abstract:

This paper examines non-response in a mobile app study designed to collect expenditure data. We invited 2,383 members of the nationally representative *Understanding Society* Innovation Panel in Great Britain to download an app to record their spending on goods and services: participants were asked to scan receipts or report spending directly in the app every day for a month. We examine coverage of mobile devices and participation in the app study at different stages of the process. We use data from the prior wave of the panel to examine the prevalence of potential barriers to participation, including access, ability and willingness to use different mobile technologies. We also examine bias in who has devices and in who participates, considering socio-demographic characteristics, financial position and financial behaviours. While the participation rate was low, drop out was also low: over 80% of participants remained in the study for the full month. The main barriers to participation were access to, and frequency of use of mobile devices, willingness to download an app for a survey, and general cooperativeness with the survey. We found extensive coverage bias in who has and does not have mobile devices, and some bias in who participates conditional on having a device. In the full sample, biases remain in who participates in terms of socio-demographic characteristics and financial behaviours. Crucially, however, we observe no biases for several key correlates of spending.

Keywords: mobile app, scanning, finances

JES Classification: D14, C81, C83

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

Introduction

The well-documented rise in the use of mobile devices brings many opportunities for survey researchers to enhance and extend measurement (see Link, et al. 2014). But using these technologies to improve survey measurement also presents challenges. Some of these are related to coverage, or differential access to or use of the technologies. With the increasing use of mobile devices, this has become more nuanced than the standard "digital divide" of the haves and have-nots. As Hargittai (2002) has termed it, the second-level digital divide distinguishes people based on how they use the technology, rather than just whether or not they have the technology. Another key source of potential selection bias in the adoption of mobile-enabled technologies is that of non-response. Non-response can occur at many stages, from consent to participate, to downloading and installing an app or device, to using that app (whether actively or passively) to capture and transmit data, often repeatedly over a period of time. In addition, the measurement properties of these new methods are not yet well understood. While there is a vast range of new possibilities, and many different ways to implement studies, there is scant research on the impact of mobile technologies on total survey error, and on the costs and efficiency of survey data collection (Jäckle, Couper, Gaia and Lessof in press). There are many unanswered questions about how best to integrate these new technologies into survey data collection.

The focus of this paper is on one particular type of app use in which participants are requested to download an app and then actively use that app to provide data: we asked members of a large-scale probability household panel to download and use an app to scan receipts for purchases, record a purchase without a receipt, or report a day without purchases over the course of a month. The specific focus of this paper is on the non-response associated with this app data collection activity, examining the rates and biases associated with both mobile device coverage and participation in the app study.

Background

As the frequency and intensity of measurement increases, so too does the complexity of the non-response problem. There are more opportunities for persons to become non-respondents, and more reasons why non-response may occur. As survey researchers start to explore the use of mobile-enabled technologies, understanding non-response and its possible effects on the data being produced becomes more important.

In cross-sectional surveys, unit non-response is often thought of as a binary outcome: a sample member either participates in the survey or they do not. Respondents may drop out (or break off) before completing the survey, or they may complete the survey without answering all items (item non-response). Increasingly in surveys, respondents may be asked to complete additional tasks, such as physical or cognitive tests, provision of biosamples, linkage to administrative records and the like (e.g. Sakshaug, Couper, Ofstedal and Weir 2012, O'Doherty, et al. 2014, Benzeval, Kumari and Jones 2016). These are often viewed as separate tasks for which consent is requested, and may produce additional sources of non-response. In longitudinal surveys, the addition of attrition and wave non-response may further increase the opportunities for non-participation.

There are a number of different barriers that may affect participation in an app-based activity and lead to selection biases in the achieved sample of those who complete the task as requested. An initial barrier is access to or use of a mobile device capable of installing apps. This is usually viewed as a problem of coverage, with differential access to devices potentially producing selection bias. Among those with suitable devices, the question is then whether respondents are able and willing to participate in such a study. Respondents are requested to download and install an app, then use that app along with related features (e.g., a camera to take pictures of receipts). Respondent familiarity with, and comfort using, various features of mobile devices are likely to play a role. Physical capacity (e.g., vision or dexterity) may also limit participation. The technical capabilities of the mobile device (e.g., storage capacity) may also affect whether the app can be successfully installed. Thus, various factors may affect the respondent's ability to complete the task. A further set of barriers relates to respondents' willingness to engage in such an activity. This may in turn be related to general willingness to participate in surveys and share personal information, as well as reactions to specific features of the requested task. Time constraints are another factor that may affect willingness to participate in a relatively burdensome app-based activity. General concerns about confidentiality and privacy issues relating to technology, as well as specific concerns about sharing personal information on spending, may also affect willingness.

The above factors are likely to affect initial agreement to participate in the study, if consent to participation is an explicit step in the process. Additional non-response during the process of downloading, installing, and registering the app (initial set-up) can occur. Once the app is working, participants must then remember to use it for each shopping event, or for each receipt received, or to report each day that no purchases were made. This requires continued motivation and engagement. The experience of participating in the study may affect ongoing compliance, in similar fashion to participation in ongoing diary studies or other studies requiring intensive measurement

(see e.g. Silberstein and Scott 1991). Participants may lose motivation or interest, leading to drop out or attrition. They may forget to scan or report certain events, leading to missed activities. Participants may choose to report certain types of spending but not others, similarly leading to differential exclusion of shopping events. Insufficient battery power, storage limitations, and other technical limitations may also lead to missed events. The focus of this paper is on initial and continued participation in the study (i.e., unit non-response) rather than missed events (item non-response).

There are thus a wide range of factors that may affect participation in an intensive app-based study such as this, and many points at which non-response may occur. Given the rising use of mobile devices for these types of research activities (whether in the fields of health, transportation, finances, or some other domain), research is needed on the causes and consequences of non-participation in mobile-based studies. Much of the existing literature focuses on small groups of volunteers. While research on the non-response issue is starting to emerge, the literature is still very sparse and few studies have examined the rates of coverage and participation at each stage and the nature of selection biases that may result.

Coverage is usually viewed as a fixed attribute of a sample unit (see Groves, et al. 2009, chapter 3), but the issue is much more fluid when considering mobile device use for specific tasks (see also Hargittai 2002, Couper, et al. 2018). A few papers have examined coverage bias of mobile devices (e.g. Fuchs and Busse 2009, Metzler and Fuchs 2014, Antoun, Conrad, Couper and West 2018, Couper, et al. 2018), while more papers have looked at socio-demographic differences in the use of mobile devices for web survey completion (e.g. Lugtig, Toepoel and Amin 2016, Brosnan, Grün and Dolnicar 2017, Maslovskaya, Durrant, Smith, Hanson and Villar 2017, Metzler and Fuchs 2017). The findings on age are generally consistent, with greater mobile access or use by younger persons (although Brosnan, et al. 2017, find more tablet use by older persons), but the results are inconsistent with respect to education and gender. Given the rapidly changing landscape with regard to mobile device penetration, the evidence on coverage bias is mixed.

Several papers have examined stated or hypothetical willingness to engage in various tasks using mobile devices. Armoogum and colleagues (2013) asked respondents about their willingness to use a GPS device in a travel survey in France. About one-third (30%) said yes without conditions, while 5% agreed as long as they could turn it off, and 64% said no. Biler, Senk, and Winkerova (2013) similarly asked respondents about GPS tracking in a travel survey in the Czech Republic: only 8% said they were willing, with 25% uncertain, and 57% not willing. Revilla and colleagues (2016) elicited

willingness to do three additional tasks among members of an online panel in several countries: 1) share GPS location, 2) install an app, and 3) take a photo. They found the level of willingness to be relatively high but varying across countries (e.g., 30% of respondents in Mexico and 17% in Portugal agreeing to share location through GPS) and across tasks (e.g., 24.2% in Spain agreeing to GPS location-sharing, 29.2% to take photos, and 35.5% to install an app). Revilla, Couper and Ochoa (2017) reported on Spanish panel members' willingness to do twenty different hypothetical tasks, including installing a passive tracking app, passive GPS tracking, and sharing photos and social media content. They again found that respondents were more willing to do some tasks than others. Stated willingness was generally higher for tasks where respondents have control over the reporting of the results than for passive tracking, even if the former requires more work on the part of respondents. Using data from the Understanding Society Innovation Panel in Great Britain, Wenz, Couper and Jäckle (in press) also found that stated willingness differed markedly between different types of tasks; that respondents were more willing to do tasks that required their active participation than tasks that collect data passively; and that they were less willing to do tasks that require downloading an app or that are potentially threatening to their privacy. Finally, Keusch et al. (2017) asked members of a German online panel about their willingness to install an app that passively tracks the usage of their smartphone. Respondents were shown vignettes with varying characteristics of the task. The results suggest that respondents would be more willing to participate in such a task if it is sponsored by a university rather than a government agency, if data are collected over a shorter period of time, if respondents have the possibility of temporarily switching off the app, if they are offered incentives, and if they were not asked to fill in questionnaires in addition to installing the app.

Some studies have analysed actual compliance with the requests to provide additional data using mobile technologies. For example, in a panel study of college students in the U.S., Crawford et al. (2013), found that 58% said yes to a hypothetical question about GPS capture. In a subsequent wave of the survey, between 20% and 33% of survey respondents (depending on the consent condition) provided usable GPS data. Toepoel and Lugtig (2014) asked members of a Dutch panel for the one-time capture of GPS coordinates. They report that 26% of smartphone participants and 24% of PC participants agreed to such capture. The LISS Mobile Mobility Panel in the Netherlands recruited panel members with smartphones to provide GPS data. Of those invited, 56% downloaded the app, activated Wi-Fi and GPS, and provided data for at least 1 day (Scherpenzeel 2017). Angrisani, Kapteyn, and Samek (2017) invited panellists of the Understanding America Study to sign up to a financial aggregator and provide access to the data collected by the aggregator to researchers. They

report that 45.8% consented to the request, 32.0% signed up with the financial aggregator, and 12.2% linked one or more financial institutions to their accounts.

A few papers have explored factors related to non-response or examined potential non-response bias (e.g. Armoogum, et al. 2013, Biler, et al. 2013, Pinter 2015, Revilla, et al. 2016, Keusch, et al. 2017, Revilla, et al. 2017, Wenz, et al. in press). With regard to socio-demographic correlates the results are somewhat mixed. For example, while Armoogum et al. (2013), Biler et al. (2013) and Revilla et al. (2016) found that younger persons were more willing to participate, Wenz et al. (in press) found no effects of age, and Revilla et al. (2017) found an effect of age only for activities over which respondents have control. Results are also mixed with respect to gender. Similarly, Armoogum et al. (2013) found that those in smaller households were more willing to use a GPS device, while Biler et al. (2013) found that those in large households were more willing. Several of the studies found that factors related to familiarity or experience are positively related to willingness, as are attitudes concerning privacy, confidentiality, and trust. These findings point to the need for further research on socio-demographic and attitudinal differences in non-response on tasks such as this.

Given the wide range of additional tasks that can be performed, very little is known about compliance with actual requests to use mobile devices for research activities. In this paper we focus on one particular activity, the installation and use of a spending app to scan and transmit receipts over a period of a month. We examine a number of different outcomes related to coverage and non-response, from having a mobile device, to downloading and installing the app (agreeing to participate in the study was not a separate step), to using it at least once, to daily participation over the month of the study. Specifically, we address the following research questions:

1. What are the mobile device coverage and app participation rates in a mobile app study of the general population?
2. Do incentives increase participation? Do survey non-respondents engage in the app study?
3. Which devices do participants use and does device choice correspond to previously stated preferences?
4. What are the patterns of participation over the month?
5. What are the main reasons that mobile device users state for not participating in the app study?
6. How prevalent are potential barriers to participating in the app study? Which are most important in predicting participation?
7. What is the nature of coverage and participation bias? Are coverage and participation related to financial behaviours and outcomes?

Data

The *Understanding Society* Innovation Panel

The Innovation Panel is part of *Understanding Society*: The UK Household Longitudinal Study (University of Essex. Institute for Social and Economic Research 2017). The general survey design mirrors that of the main *Understanding Society* study (https://www.understandingsociety.ac.uk), with the difference being that the primary purpose of the Innovation Panel is methodological testing and experimentation (see Jäckle, Gaia, Al Baghal, Burton and Lynn 2017 for further details). The Innovation Panel is a clustered and stratified sample of 1,500 households in Great Britain that have been interviewed annually since 2008. All household members aged 16+ are interviewed about their socio-economic circumstances, health and family situation and other rotating topics. One person completes an additional household questionnaire about the conditions, tenure, and costs of their housing. Individuals are followed if they move within the country. Refreshment samples of approximately 500 participant households were added at waves 4 and 7. In this paper we use data collected in wave 9 (IP9) as predictors of participation in the spending study. Fieldwork for IP9 took place between May and September 2016. The IP9 household response rate was 84.7% (Jäckle, et al. 2017). Sample members in a random two-thirds of households were invited to complete the survey online, and if they did not respond within two weeks they were followed up by face-to-face interviewers. The remaining third of the sample were issued to face-to-face first. Both samples included a final mop-up stage in which non-respondents were followed up by telephone and web. The Innovation Panel data are available from the UK Data Service at https://discover.ukdataservice.ac.uk/catalogue/?sn=6849.

The spending study

All adult sample members in households where at least one person gave an interview in IP9 were invited to participate in the spending study (n=2,383).[1] The analyses presented in this paper are restricted to those who gave a full interview in IP9 and were invited to the spending study (n=2,112). The study was carried out in collaboration with Kantar Worldpanel, who developed the app and implemented fieldwork between the end of October 2016 and early January 2017 (University of Essex. Institute for Social and Economic Research 2016). Each sample member was sent a letter inviting them to download the app to their smartphone or tablet and to use it to report purchases of

---

[1] The data and documentation from the spending study will be available from the UK Data Service. Until then the documentation is available at https://www.iser.essex.ac.uk/research/projects/understanding-household-finance-through-better-measurement.

goods and services for a month. The app was compatible with iOS and Android operating systems. Note that all adults in IP9 respondent households were invited to participate in the app study, regardless of whether they had internet access or a suitable mobile device. Sample members for whom an email address was known also received the invitation by email. The letter contained a unique log-in to a registration survey, as well as the rationale for the study, information about incentives and a Frequently Asked Questions (FAQ) section on the back page. There was also a link to a more extensive FAQ section online, which was updated as the study went along (see Appendix). Reminders were sent twice a week by email for three weeks to anyone who had not yet completed the registration survey, and a final reminder letter was sent by post in the fourth week. In the app, the participant could scan and upload a receipt, record a purchase without a receipt, report a day without purchases, and also access the FAQs. The app also sent push notifications at around 5pm each day to remind people to scan any receipts they had.

In the advance letter, sample members were told that they would earn either £2 or £6 for downloading the app (households were randomly allocated to groups), plus £0.50 for every day on which they used the app, plus a £10 bonus at the end of the study if they used the app every day, plus £3 if they completed a short end of project questionnaire. The maximum incentive participants could earn was either £30.50 or £34.50, depending on the experimental group they were assigned to. Participants received their reward by post after completing fieldwork, in the form of a gift voucher that can be used in many high street shops.

The unique link sent in their advance letter led participants to a short registration survey designed to verify their identity, collect their email address, and ask a few short questions about their purchasing behaviours. At the end of the registration survey each participant was given their unique app ID, instructions on how to download the app, and was sent an email acknowledgement which included their unique app ID code and links to the app on the two main app stores (App Store and Google Play). At the end of each week in which respondents used the app at least once, they were sent an email confirming how much they had earned that week and their reward balance, and asking them to complete a short end of week survey about their experiences with and use of the app that week (data not used in this paper). At the end of fieldwork all sample members were emailed a link to an online end of project questionnaire, with questions tailored to participants who had completed the full month, participants who dropped out before the end of the month, and non-participants. Non-respondents to this online survey were sent a paper questionnaire by post, with a Freepost return envelope, but no incentive. The response rate for the end of project survey was 88.9% for those who used the app at least once and 33.6% for non-participants.

Outcomes: measures of coverage and participation

The indicator of mobile device coverage was derived from a question in the IP9 interview[2] asking *"Which of the following devices do you use to connect to the Internet? [Desktop computer, laptop, smartphone, tablet, feature phone/non-touchscreen mobile phone, E-book reader (e.g. Kindle), Smartwatch, other]"*. Respondents were coded as having a mobile device if they reported using a smartphone, tablet, or both. They were coded as not having a mobile device if they mentioned neither a smartphone nor a tablet, or if they had indicated that they do not use the internet for personal use.

The measures of participation are derived from the app paradata which recorded a total of 11,507 app uses from the 270 participants who used the app at least once. The paradata recorded the start time of each activity, the end time when the data or scanned image finished uploading, the device used, and the activity type: whether the app was used to scan a receipt, record a purchase, or report a day without purchases. The outcomes examined in the analyses are:

- **Completed registration survey:** coded as 1 if the sample member completed the registration survey, and 0 otherwise.
- **Used app at least once:** coded as 1 if the app paradata contain at least one observation on the sample member, and 0 otherwise.
- **Used app for five weeks:** based on the recorded start time this outcome is coded as 1 if the paradata contain at least one observation on the sample member in each of five consecutive calendar weeks, and 0 otherwise.
- **Device used:** derived from the agent user string and coded as either smartphone or tablet.
- **Daily app use:** derived from the start time and activity. Further explanations in the text relating to Figures 1 and 2.

Predictors of participation

All variables measuring potential barriers to participating in the app study are from the IP9 interview. In the face-to-face interviews these questions were asked using Computer Assisted Self-Interviewing (CASI), for which the interviewer handed their laptop over to the respondent. In the web version, all questions were in the same order as in the face-to-face interview; the self-completion section was not distinguished from other modules in the questionnaire. The analysis of predictors of participation is conditional on having a mobile device.

---

[2] The IP9 questionnaire can be found at https://www.understandingsociety.ac.uk/documentation/innovation-panel/questionnaires.

We group the predictors of participation into four related sets of variables: access to mobile technologies, ability to use such technologies, willingness to use them, and general survey cooperativeness. The variables related to *access* to mobile technologies include:

- **Frequency of internet use:** how often the respondent uses the internet for personal use, coded as every day, several times a week or less frequently.

- **Type of mobile device:** derived from the question asking *"Which of the following devices do you use to connect to the Internet? [Desktop computer, laptop, smartphone, tablet, feature phone/non-touchscreen mobile phone, E-book reader (e.g. Kindle), Smartwatch, other]"*. Two indicators coded as 1 if the respondent has a smartphone/tablet, and 2 if not.

- **Wi-Fi at home:** coded as 1 if respondent has Wi-Fi access at home, and 2 if not.

- **Data plan:** coded as 1 if the respondent has a fixed data plan to get mobile internet on their smartphone, 2 if they have a pay-as-you-go contract, and 3 if they have neither.

The variables related to *ability* to use mobile technologies are derived from questions about the respondents' usage of their mobile devices. For concepts where we asked the same question separately about smartphones and tablets, the question text documented below refers to "[smartphone/tablet]" to avoid repetition. For respondents who have both a smartphone and a tablet the variables are coded as the higher of the scores for the two devices.

- **Frequency of device use:** derived from the questions "*How often do you use a smartphone for activities other than phone calls or text messaging?"* and *"How often do you use a tablet?"* Coded as 1 if respondent uses at least one of the devices every day, and 2 if less often.

- **Self-reported skill:** Derived from the two questions *"Generally, how would you rate your skills of using a [smartphone/tablet] on a scale from 1 = Beginner to 5 = Advanced?"* coded as advanced if categories 4 and 5 for either device, medium if categories 2 and 3, and beginner if category 1.

- **Takes photos, online purchases, online banking, installs apps:** based on questions asking for which activities respondents use their smartphone and/or tablet. Each variable is coded as 1 if the respondent does the activity on at least one of their devices, and 2 if not.

The variables related to *willingness* to use mobile technologies include:

- **Willingness to download app, willingness to use camera:** derived from questions asking *"How willing would you be to carry out the following tasks on your [smartphone/tablet] for a survey?"* The activities asked about included *"Download a survey app to complete an online questionnaire"* and *"Use the camera of your [smartphone/tablet] to take photos or scan*

barcodes". Coded as 1 if very or somewhat willing on at least one device, and 2 if a little or not willing.

- **Security concerns: complete online via app, use camera for barcodes:** derived from questions asking "*In general, how concerned would you be about the security of providing information in the following ways?*" The data collection methods asked about included "*Download a survey app to complete an online questionnaire*", and "*Use the camera on your [smartphone/tablet] to take photos or scan barcodes*". Coded as 1 if not at all concerned on at least one of their devices, 2 if a little or somewhat concerned, and 3 if very or extremely concerned.

Additional variables related to *general cooperativeness* with the survey and willingness to share personal information include:

- **Item non-response rate:** the proportion of eligible questions in the IP9 individual interview to which the respondent answered "don't know", "refused" or that were otherwise missing. The base excludes ten questions about receipt of State welfare and pensions, which are repeated for each income source reported.
- **Consent to data linkage:** coded as 1 if the respondent gave consent in IP9 for their survey data to be linked to credit rating data about them held by the Financial Conduct Authority, and 0 otherwise.
- **Mode of interview:** coded as face-to-face versus web.

The combined item non-response rate for predictors of participation due to "don't know" and "refused" responses were mostly ≤0.5% (21 items), below 1.2% for a further 8 items, and 2.1% for the consent to linkage question. Due to the low rates of missingness, we set missing cases for a given variable to the category mentioned last in the descriptions above.

All variables used for the analysis of coverage and participation bias are also from the IP9 interview:

- **Socio-demographic characteristics:** gender, age and highest educational qualification.
- **Financial position:** these are outcome variables that we expect to be correlated with the monthly expenditure measured by the app and that were asked of the full sample in either the individual questionnaire or the household questionnaire. Variables from the individual questionnaire are the respondent's personal monthly income, derived from the sum of all reported income sources, and their subjective assessment of how well they are getting by financially. Variables derived from the household questionnaire are household expenditure on food (groceries plus food consumed outside the home) in the last month, household

expenditure on fuel (gas, electricity, oil or other) in the last year, and whether the household is behind or struggling with any payments for housing costs or utility bills.

- **Financial behaviours:** whether and how the respondent keeps a budget, how often they check their bank balance, how they check their balance, whether they file a tax return, and which (if any) store loyalty cards they have.

For household spending on food and fuel we treat missing observations as a separate category (see Table 5). For all other items the combined percentage of "don't know" and "refused" responses was ≤ 1.5%. Due to the low rates of missingness we use case-wise deletion and include only respondents with non-missing observations in testing for bias in those variables. Some additional variables used in the analyses are described in context in the Results section.

All standard errors account for the clustered and stratified sample design of the Innovation Panel.

Results

RQ1: What are the mobile device coverage and app participation rates in a mobile app study of the general population?

Among all IP9 respondents 16.5% completed the registration survey and 12.8% used the app at least once. This is very similar to the participation rate of 12.2% reported by Angrisani et al. (2017). Subsequent drop-out was unexpectedly low: 10.2% of IP9 respondents used the app at least once in each of the five consecutive weeks (Table 1).

Not everyone in the IP9 respondent sample however had a mobile device (Table 1): 76.3% of respondents reported using a smartphone or tablet, 20.7% reported not using a smartphone, tablet, or the internet, and a further 3.0% did not answer the questions about mobile device use. This latter group includes 24 CAPI respondents who declined to do the self-completion section, 26 CATI respondents who were not asked the self-completion section by design, and 13 respondents who completed the self-completion section but did not answer the question about mobile devices. Among IP9 respondents who reported having a mobile device completion rates were somewhat higher than in the full sample: 20.2% of mobile device users completed the registration survey, 15.8% used the app at least once, and 12.8% used the app at least once in each of the five weeks.

As the numbers in Table 1 indicate, there were 15 respondents who did not report having a mobile device in the IP9 interview, but who nonetheless used the app at least once. This group included 12 respondents who reported not using the internet or not having a smartphone or tablet in the IP9

interview, two respondents who did not complete the self-completion section, and one respondent who did not answer the mobile device question. We cannot identify whether these respondents mis-reported their device usage in the IP9 interview, whether they purchased devices in the months between their interview and the Spending Study, or whether they used someone else's device in order to participate.

Depending on the research question, the analyses that follow are based on different sub-samples: the full IP9 respondent sample, mobile device users, or participants in the Spending Study. Which sample is used is documented in the results section for each research question.

Table 1: Participation in the spending study

	Full sample		Mobile users	
	N	%	N	%
Issued sample (IP9 respondents)	**2,112**	**100.0**		--
Did not answer mobile device questions	63	3.0		--
Did not report having a mobile device at IP9	438	20.7		--
Had a mobile device in IP9 interview	**1,611**	**76.3**	**1,611**	**100.0**
Completed the registration survey	348	16.5	326	20.2
Used app at least once	**270**	**12.8**	**255**	**15.8**
Used app at least once in each of five weeks	216	10.2	206	12.8

RQ2: Do incentives increase participation? Do survey non-respondents engage in the app study?

The incentive experiment, varying the value of the bonus for downloading the app, had no effect on participation outcomes: the proportion of IP9 respondents who completed the registration survey was 15.9% in the £2 group and 17.0% in the £6 group; the proportion who used the app at least once was 11.9% and 13.6% respectively, and the proportion who used the app at least once in each of the five weeks was 9.4% and 10.2% respectively. In Chi2 tests adjusted for clustering and stratification, none of the differences in outcomes between the £2 and £6 treatment groups were significant at the 10%-level.

We also invited IP9 non-respondents, living in households with at least one IP9 respondent, to participate in the app study (n=271). This was to test whether people who do not participate in the annual survey interview might be interested in participating in data collection activities using other technologies. However only 2.2% completed the registration survey and 1.5% used the app at least once. As we have no data on the covariates collected in the IP9 interviews for this sub-sample, we exclude IP9 non-respondents from further analyses in this paper.

RQ3: Which devices do participants use and does device choice correspond to previously stated preferences?

Among the 270 participants who used the app at least once, the majority used smartphones, regardless of hypothetical preferences stated in the previous interview. According to the app paradata, 82.6% of participants used the app on a smartphone, 15.6% used a tablet and 1.9% used both types of devices. For participants who reported having both devices in the IP9 interview (n=182), Table 2 shows which device they used, by how willing they said they would be to use the camera of their smartphone/tablet to take photos or scan barcodes for a survey. Even among respondents who had indicated a greater willingness to use their tablet for this purpose, 62.5% actually used their smartphone, as did 75.0% of participants who had said they would not be willing to use either device. Everyone who reported higher willingness to use their smartphone acted according to their stated preference and used a smartphone.

Table 2: Device used by hypothetical willingness (participants with both devices, row %)

Hypothetical willingness	Used smartphone	Used tablet	Used both	N
Equally willing on both devices	86.5	11.5	2.1	96
More willing on smartphone	100.0	0.0	0.0	50
More willing on tablet	62.5	31.3	6.3	16
Not willing on either device	75.0	25.0	0.0	20
Total	86.8	11.5	1.7	182

RQ4: What are the patterns of participation over the month?

The solid line in Figure 1 shows the daily participation rates among the 270 participants who used the app at least once, starting with the day on which they first used it. App use includes scanning receipts, entering spending information without a receipt, or declaring no purchases for that day. On day 2 only 75.9% of participants used the app. From day 2 onwards the drop-out rate was much lower than expected, with 60.7% of participants still using the app on day 31. The solid line however hides the non-monotonic nature of drop-out: respondents who missed a day tended to continue using the app on a future day. The dashed line in Figure 1 shows for each day, the proportion of participants who continued to use the app on at least one day in the future. The area above the dashed line therefore represents permanent drop-out. Only 4.8% of participants did not use the app again after the first day and a striking 81.5% remained in the study for at least 29 days. Anecdotal feedback from participants suggests that the £10 bonus promised if they used the app every day for the entire month was a strong motivator.

Figure 2 shows the mean number of times participants scanned a receipt or reported a purchase in the app, for each of the 31 days. That is, unlike Figure 1, this graph excludes app uses to report no purchases for the day. The graph distinguishes participants by how often they had reported spending money in the registration survey, where they were asked: *"How often do you spend money on goods or services? [Several times a day, about once a day, more than twice a week, once or twice a week, less than once a week, never]";* 11.3% reported spending money several times a day, 27.8% about once a day, and 60.9% less than once a day. A small number (n=4) of respondents answered "don't know" or "refused" and are excluded from Figure 2. Those who said they spend money more than once a day scanned receipts or reported purchases on average 1.2 times per day. This was significantly higher than those who spend money less than once a day: they scanned or reported purchases on average 0.8 times per day (P = 0.018). Those who reported spending about once a day scanned or reported purchases on average 0.9 times per day (P > 0.05 for both comparisons with the other groups).

The average number of app uses varies somewhat across the 31 days (Figure 2). On day 1 there is a clear difference in the means between the three groups: those who reported spending money more than once a day used the app to scan receipts or report purchases on average 2.7 times, those who spend about once a day used the app on average 1.6 times, and those who spend less frequently used it on average 1.2 times. The 95% confidence intervals of the daily means for the three groups overlap, with two exceptions for the groups with the lowest and the highest spending frequency: on day 1 (mean 1.2, C.I. 0.93, 1.39 versus mean 2.7, C.I. 1.47, 3.93) and on day 20 (mean 0.7, C.I. 0.54, 0.88 versus mean 1.8, C.I. 1.06, 2.54). The confidence intervals are not shown in the graph to maintain readability. Although the daily means fluctuate, it is striking that the number of times participants scan receipts or report purchases is stable until day 31.

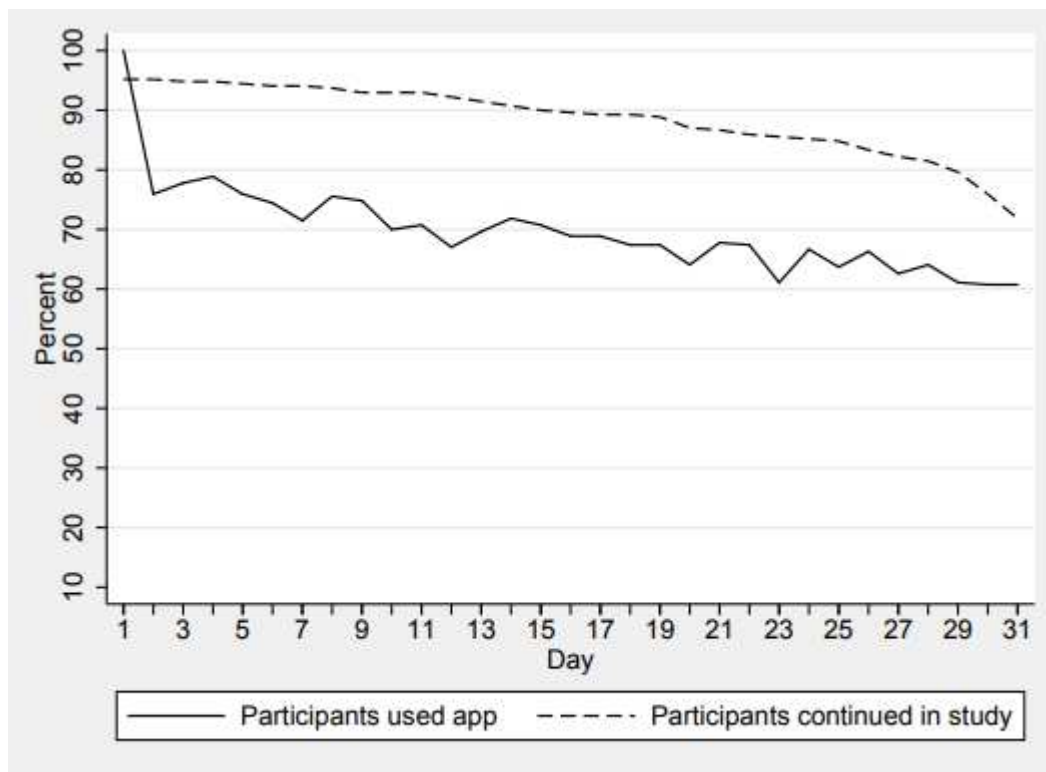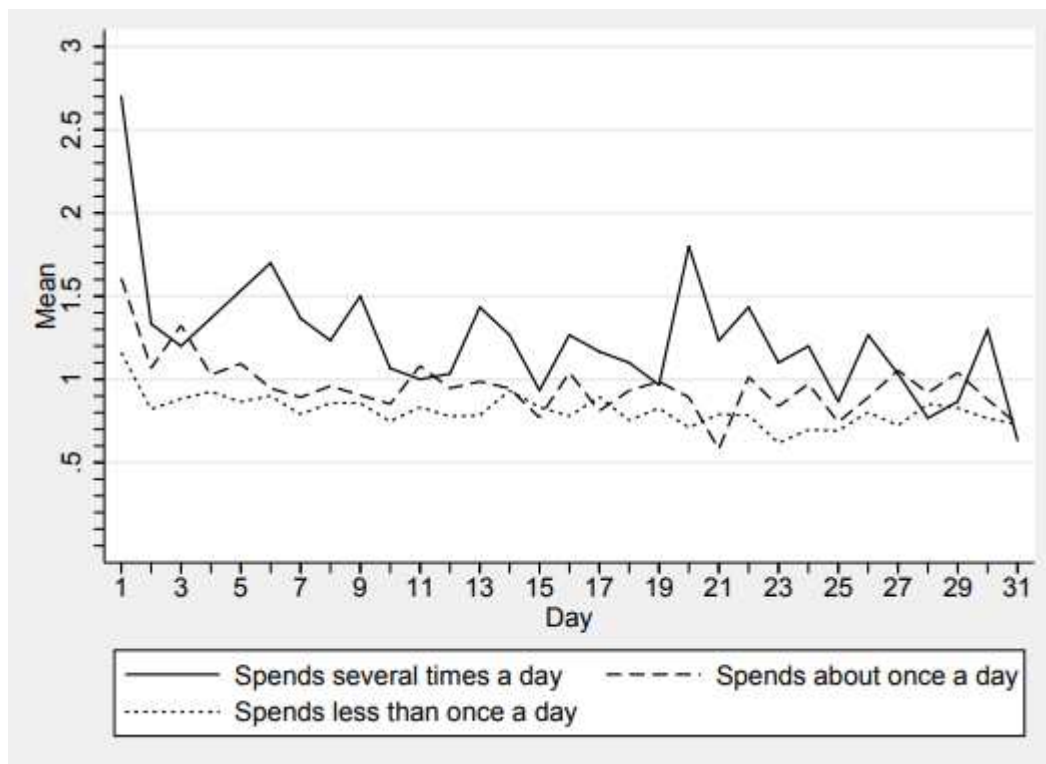**Figure 1: Percent of app users and drop-out per day**

Figure 2: Mean number of receipt scans and purchases entered, by self-reported frequency of spending money on goods and services

There are several possible explanations for the steep drop-off in the mean number of scans and reported purchases after day 1. The first time respondents used the app they were more likely to scan receipts that were a few days old (see Lessof, Jäckle and Couper 2017).[3] This would account for a larger number of receipts scanned on the first day compared to later days. In addition respondents might have learnt that they only needed to use the app once each day to get their daily reward of £0.50. However, given that we did not experiment with the daily incentive, we cannot rule out other explanations.

RQ5: What are the main reasons that mobile device users state for not participating in the app study?

To examine the reasons why those who could in principle participate in the app study do not, we focus on IP9 mobile device users (n=1,611), who did not participate in the app study (n=1,356), and did not complete the registration survey for the app (n=1,281)[4], but did complete the end of project debrief survey (n=425).

Table 3 shows the responses given to two check-all-that-apply questions in the end of project survey: *"When deciding whether to participate in the spending study, which of the following difficulties did you have?", "And which of the following applied to you?"* These were effectively a single question, split into two because of the large number of response options. Of the 425 non-participants who answered the end of project survey, 348 reported at least one difficulty. The most frequent single response was that respondents did not have time to scan (39.6%). Just over half (53.5%) of respondents mentioned one or more technical problems: they did not have a smartphone or tablet which can download apps, the storage space on their device was insufficient to download the app, the app was not compatible with their operating system, they could not find the app in the app store or the link to downloading the app did not work. However without knowing details of the devices used, we cannot distinguish genuine technical problems from user errors. Nearly half (46.5%) mentioned at least one privacy concern: they were not willing to share spending information, or not confident that information would be held securely. Finally, 41.6% mentioned lack of confidence, either with using their mobile device for this kind of activity and/or with downloading apps, and 11.1% said they were not interested.

---

[3] Receipts where the date on the receipt preceded the day on which the invitation to the spending study was sent out (n=34) were dropped from Figure 1 and Figure 2, although results are unchanged if they are included.
[4] Non-participants who had completed the registration survey were routed into a different question in the debrief questionnaire.

Table 3: Reasons for not participating in the app study

	N	% of cases
Did not have time to scan	168	39.6
Did not try to download the app	126	29.7
Not willing to share spending information	84	19.8
Not confident using my phone or tablet for this kind of activity	75	17.7
Not able or confident to download apps onto my phone or tablet	66	15.6
Do not have a smartphone or tablet which can download apps	60	14.2
Not confident that information would be held securely	60	14.2
Not interested	47	11.1
Did not have sufficient storage space to download the app	40	9.4
Do not have access to the internet on my phone or tablet	23	5.4
Could not download the app because not compatible with operating system	18	4.3
Link to downloading the app did not work	13	3.1
Could not find the app in the app store	8	1.9

Notes: n=425. Multiple mentions.

RQ6: How prevalent are potential barriers to participating in the app study? Which are most important in predicting participation?

Table 1 above revealed that device ownership remains an important barrier to participation in mobile data collection tasks: only 76.3% of IP9 respondents reported using a smartphone or tablet to connect to the internet. As a comparison, the Ofcom statistics for 2017 Q1 indicate that 76% of adults in the UK had a smartphone (Ofcom 2017).

The following analysis focuses on barriers for respondents who did have a mobile device at the time of the IP9 interview (n=1,611). Table 4 shows the prevalence of different potential barriers relating to access, ability and hypothetical willingness to use the spending study app. Column 1 shows that *access* is not much of a barrier among mobile device users: only 4.2% used the internet less than several times a week, with most (84.9%) using it daily, and nearly all had Wi-Fi in their home (97.6%). *Ability* to participate in an app study was similarly high: 83.6% used at least one of their devices daily, 95.8% considered themselves advanced or intermediate users, and between 59.0% and 88.4% used at least one of their devices to take photos, make online purchases, use online banking or install apps. *Willingness* however seems to be more of a barrier: only half (51.1%) said they would be very or somewhat willing to download an app and 62.3% to use the camera on either device for a survey. Only a quarter (25.7%) would not at all be concerned about the security of providing information by downloading an app to complete an online questionnaire and one fifth (19.7%) would not be concerned about using the camera on their device to take photos or scan barcodes. Willingness might also depend on more *general cooperativeness* with the survey and willingness to

share personal data, for which item non-response and consent to data linkage from the IP9 interview are used as indicators. The item non-response rate among mobile device users ranged from 2.0% to 29.3%, with a median of 3.9%. Consent to data linkage was given by 58.3% of mobile device users. As consent was lower among respondents who completed their questionnaire online, we control for the mode of interview in the regression models: 58.0% of mobile device users completed their interview online, 42.0% completed in a face-to-face interview.

The bivariate relationships between each of the potential barriers and whether a sample member used the app at least once are strong (Column 2): for each of the potential barriers the Chi2 test is significant at $P < 0.05$ or less. There are two exceptions: whether or not the respondent has a tablet, and whether or not they have Wi-Fi at home are not related to the probability of using the app. According to the bivariate tests the strongest predictors of participation appear to be advanced self-reported skill using their mobile device (20.2% participated), using at least one device for online banking (20.0%), being very or somewhat willing to download an app for a survey (21.5%), and being not at all concerned about the security of providing information by downloading an app to complete an online questionnaire (23.1%).

Columns 3 to 6 show the average marginal effects estimated from probit models of the probability of using the app at least once. Column 3 shows the results of four separate models, including in turn the predictors relating to (1) access, (2) ability, (3) willingness, and (4) general cooperativeness. Column 5 shows the results of the full model. The Hosmer–Lemeshow goodness-of-fit test (Archer and Lemeshow 2006), which can be used for logit or probit regression models taking survey design into account, suggests good model fit with Prob>F ranging from 0.548 to 0.999 for each of the four partial models and the full model.

Of the predictors related to *access*, using the internet every day increased the probability of participating by 11.4 percentage points, compared to only using it several times a month or less (Column 3). The joint test of the overall effect of frequency of internet use is significant with Prob>F = 0.005 (Column 4). Having a smartphone and having a tablet increase the probability by 8.8 and 5.7 percentage points respectively. Of the predictors relating to *ability*, using the device every day increased the probability of participating by 6.1 percentage points, using at least one device for online banking increases it by 4.8 percentage points, and installing apps by 5.7 percentage points. Self-rated skill is also a significant predictor according to the joint F-test of whether both coefficients equal zero (Prob>F = 0.022). Of the predictors related to *willingness*, being very or somewhat willing to download an app for a survey increase the probability of participation by 8.8 percentage points.

Finally, the indictors of *general cooperativeness* with the survey each increased the probability of participating by between 5.4 and 8.4 percentage points.

In the full model (Column 5) the only predictors that remain significant are using one of the devices daily (+6.5 percentage points) and being very or somewhat wiling to download an app for a survey (+5.1 percentage points). All of the general cooperativeness indicators remain significant although the effect sizes are smaller than the estimates from the partial models.

Controlling for socio-demographics in the partial and full models leads to small shifts in significance levels, but the general conclusions remain largely unchanged. In a model with only gender, age (coded as 16-30, then 10 year age bands up to 70, then 71 and older), and education (coded as degree, school or other higher qualification, and lower or no qualification) predicting the probability of participation, women are more likely to participate than men (+3.1 percentage points, P=0.024), the probability of participating decreases monotonically with age (Prob>F < 0.004), but qualifications have no effect (not shown). Adding age, gender and qualifications to the models in Table 4, gender remains significant and similar in magnitude in all models except for the partial model of willingness predictors, while age is only significant in the partial model with predictors related to general cooperativeness (not shown).

Table 4: Prevalence of barriers, bivariate relationship with participation and Average Marginal Effects

		%	App used %	Partial model (AME)	Prob>F	Full model (AME)	Prob>F
Access							
Frequency of internet use	every day	84.9	17.4	0.114**		0.050	
	several times a week	11.0	7.9	0.032		0.025	
	several times a month/less	4.2	4.5***	.	0.005	.	0.669
Has a smartphone	yes	82.8	17.3	0.088*		0.043	
	no	17.2	8.7**	.		.	
Has a tablet	yes	76.2	16.9	0.057**		0.039	
	no	23.8	12.5	.		.	
Wi-Fi at home	yes	97.6	16.0	0.055		0.040	
	no	2.4	7.7	.		.	
Data plan (smartphone)	fixed data plan	69.7	17.1	-0.015		-0.043	
	pay-as-you-go contract	8.9	18.9	0.020		0.020	
	neither	21.4	10.4*	.	0.496	.	0.120
Ability							
Frequency of device use	every day	83.6	17.8	0.061*		0.065*	
	less often	16.4	6.0***	.		.	
Self-reported skill	advanced	58.7	20.2	0.064		0.010	
	medium	37.1	10.2	0.010		-0.025	
	beginner	4.2	4.4***	.	0.022	.	0.153
Takes photos	yes	88.4	17.0	0.031		0.028	
	no	11.6	7.0**	.		.	
Online purchases	yes	69.8	18.3	-0.002		-0.023	
	no	30.2	10.1***	.		.	
Online banking	yes	59.0	20.0	0.048*		0.033	
	no	41.0	9.8***	.		.	

				AME		AME	
Installs apps	yes	71.3	19.1	0.057**		0.040	
	no	28.7	7.6***	.		.	
Willingness							
Willingness to download app	very/somewhat willing	51.5	21.5	0.088***		0.051*	
	a little/not willing	48.5	9.8***	.		.	
Willingness to use camera	very/somewhat willing	62.3	18.3	0.000		-0.026	
	a little/not willing	37.7	11.7**	.		.	
Security concerns: complete online via app	not at all concerned	20.7	23.1	0.072*		0.052	
	a little/somewhat concerned	53.6	16.6	0.041		0.034	
	very/extremely concerned	25.7	8.5***	.	0.058	.	0.216
Security concerns: use camera for barcodes	not at all concerned	34.3	19.9	0.037		0.017	
	a little/somewhat concerned	45.9	15.8	0.026		0.012	
	very/extremely concerned	19.7	8.8***	.	0.439	.	0.834
General cooperativeness							
Item non-response rate > median	low item nom-response	50.0	19.5	0.054**		0.043*	
	high item non-response	50.0	12.2***	.		.	
Consent to data linkage	yes	58.3	19.2	0.084***		0.070***	
	no/don't know/refused	41.7	11.2***	.		.	
Mode of IP9 interview	face-to-face	42.0	13.0	0.066**		0.060**	
	web	58.0	17.9*	.		.	

Notes: n=1,611. AME = average marginal effects estimated from probit models of probability of using app at least once. Standard errors adjusted for clustering and stratification. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

RQ7: What is the nature of coverage and participation bias? Are coverage and participation related to financial behaviours and outcomes?

Table 5 examines the extent and nature of coverage bias due to sample members not having a mobile device, participation bias conditional on having a device, and total participation bias resulting from both non-coverage and non-participation. We test for biases in socio-demographic characteristics (gender, age, education), financial outcomes likely to correlate with the spending recorded in the app (personal monthly income, household spending, whether the household is struggling with the payment of housing costs or bills, and subjective assessments of how well the person is getting by financially), and financial behaviours (whether and how the person keeps a budget, how often and how they check their bank balance, whether they filed a tax return, and whether they have store loyalty cards).

Columns 1 and 2 in Table 5 show the distribution of each characteristic among all IP9 respondents for whom we know whether or not they use a smartphone or tablet to connect to the internet. All analyses in Table 5 exclude 63 cases for whom the device status is unknown (see the description of Table 1), resulting in an analysis sample of 2,049 respondents with known device status.

Coverage bias is documented in column 3: this shows the percentage point difference between mobile device users and the full sample with known device status. For example, 15.8% of the full sample were aged 16-30, whereas 19.3% of mobile device users were in that age range, a difference of 3.5 percentage points. Column 4 shows the P-values from Chi2 tests of the difference in characteristics between device users and non-users, adjusted for the clustered and stratified sample design. For the purposes of Table 5, the 12 respondents who had reported in IP9 that they never use the internet or do not use a mobile device, but who did then participate in the Spending Study, are recoded as having a mobile device. This results in a sample of 1,623 respondents with a mobile device.

Participation bias conditional on coverage is documented in column 5: this shows the percentage point difference between those who used the app at least once and the sample of mobile device users. For example, in the participant sample the proportion of women was 4.9 percentage points higher than among all device users. Column 6 shows the P-values from Chi2 tests of the difference between participants and non-participants, conditional on having a device. Since there were three participants for whom device status is unknown (see description of Table 1), the number of participants included in the analyses in Table 5 is 267.

Total participation bias is documented in column 7: this shows the percentage point difference between participants and the full sample for whom device status is known. The P-values in the final column are from Chi2 tests of the difference between participants and all non-participants.

The results show extensive coverage bias. For all characteristics tested, other than gender and whether the household is behind in paying bills, there are significant differences between mobile device users and non-users: mobile device users under-represent people aged 61 or older (by -10.2 percentage points), people with lower educational qualifications (by -6.1 percentage points), and those in the lower income and spending quartiles. Coverage bias is also related to financial behaviours: mobile device users are more likely than non-users to keep a budget using a computer document or spreadsheet (+2.6 percentage points), check their bank balance at least once a week (+6.0 percentage points), check their balance online (+6.7 percentage points) or using an app on a mobile device (+4.8 percentage points), less likely to check their balance using a paper statement (-6.4 percentage points), and less likely to have no store loyalty cards (-1.9 percentage points).

Conditional on coverage, there is comparatively less participation bias: among all device users there are no differences between participants and non-participants in their level of education, monthly income and spending, whether they are behind paying bills, or their subjective assessment of how well they are doing financially. There are however some differences that mirror the coverage bias: younger age groups are over-represented and older ones under-represented among participants, those who do not keep a budget, those who never check their bank balance, those who check their balance using a cashpoint or a paper statement, and those who do not have any store loyalty cards are under-represented, while those who use an app on a mobile device to check their bank balance are over-represented among participants. In addition, there are some characteristics that are related to participation, but not related to coverage: women are over-represented in the participant sample (by +4.9 percentage points), as are those who keep a budget using personal budget software on a computer or laptop (+1.8 percentage points).

Total participation bias reflects the combined effects of non-coverage and non-participation conditional on coverage. Since the non-participation rate (83.5% of device users) was so much higher than the non-coverage rate (20.8% of the sample with known device status), total bias is dominated by participation rather than coverage bias. Although there are significant differences in the financial outcome variables between device users and non-users, these disappear when examining total bias: in the full sample there are no differences between participants and non-

participants in income and spending, whether behind paying bills, and subjective assessment of how well they are doing financially.

Where there are differences between participants and non-participants in the full sample, the underlying biases related to coverage and conditional participation are in the same direction, reinforcing each other. For example, those aged 61 and over are under-represented among device users by -10.2 percentage points, and among participants conditional on device usage by -7.9 percentage points, resulting in a total non-participation bias of -18.2 percentage points. Similarly, those who keep a budget using an app on a mobile device are over-represented among device users by +4.8 percentage points and among participants conditional on device usage by +15.1 percentage points, resulting in a total bias of + 20.0 percentage points.

Overall, participation is related to socio-demographic characteristics and to financial behaviours, but not to financial outcomes. Women are over-represented among participants by +5.4 percentage points, those aged 50 or younger over-represented by +20.3 percentage points, and those with a degree over-represented by +7.5 percentage points. Those who do not keep a budget are under-represented by -9.6 percentage points, while those who check their bank balance online are over-represented by +10 percentage points, and those who use an app on a mobile device by +20.0 percentage points.

Table 5: Coverage and participation bias

	Device known		Device users		Participation/device users		Participation/known status	
[3]	N	Col %	Difference[1]	P-value[1]	Difference[2]	P-value[2]	Difference[3]	P-value[3]
Male	924	45.1	-0.5	.	-4.9	.	-5.4	.
Female	1125	54.9	0.5	0.270	4.9	0.023	5.4	0.016
16-30	324	15.8	3.5	.	2.8	.	6.3	.
31-40	267	13.0	3.1	.	6.0	.	9.1	.
41-50	360	17.6	3.1	.	1.8	.	4.9	.
51-60	426	20.8	0.5	.	-2.6	.	-2.1	.
61-70	365	17.8	-2.7	.	-5.7	.	-8.5	.
71+	307	15.0	-7.5	0.000	-2.2	0.009	-9.7	0.000
degree	506	24.7	3.1	.	4.4	.	7.5	.
GCSE, A-level, other higher	1236	60.3	3.0	.	-4.2	.	-1.1	.
other, none or missing	307	15.0	-6.1	0.000	-0.3	0.209	-6.4	0.001
personal monthly income - quartile 1	510	24.9	-1.2	.	-1.9	.	-3.2	.
quartile 2	506	24.7	-1.8	.	2.6	.	0.8	.
quartile 3	514	25.1	0.3	.	-1.0	.	-0.7	.
quartile 4	519	25.3	2.8	0.000	0.4	0.682	3.1	0.502
HH monthly spend on food - quartile 1	493	24.1	-4.6	.	0.8	.	-3.8	.
quartile 2	461	22.5	0.0	.	0.4	.	0.3	.
quartile 3	474	23.1	1.9	.	-1.5	.	0.5	.
quartile 4	482	23.5	2.6	.	2.3	.	4.9	.
missing	139	6.8	0.1	0.000	-2.0	0.640	-1.9	0.228
HH monthly spend on fuel - quartile 1	459	22.4	-1.3	.	6.3	.	4.9	.
quartile 2	572	27.9	-0.2	.	-0.4	.	-0.6	.
quartile 3	338	16.5	-0.2	.	0.2	.	0.0	.
quartile 4	460	22.4	1.1	.	-7.1	.	-6.0	.
missing	220	10.7	0.5	0.035	1.1	0.056	1.6	0.179

	N	%						
behind/struggling to pay housing costs/bills	245	12.1	0.3	0.546	-1.4	0.558	-1.1	0.637
living comfortably	629	30.8	-1.4	.	-1.0	.	-2.4	.
doing alright	843	41.3	2.2	.	1.0	.	3.2	.
just getting by or finding it difficult	568	27.8	-0.9	0.005	0.0	0.926	-0.9	0.564
keeps budget on paper (check all that apply)	521	25.7	-1.2	0.010	2.9	0.272	1.7	0.535
on computer document or spreadsheet	301	14.9	2.6	0.000	4.0	0.109	6.6	0.006
personal budget software on computer/laptop	22	1.1	0.2	0.148	1.8	0.006	1.9	0.003
online budget programme	5	0.2	0.0	0.970	0.1	0.652	0.1	0.651
personal budget app	27	1.3	0.2	0.104	0.7	0.360	0.9	0.216
do not keep a budget	1200	59.2	-1.3	0.025	-8.3	0.020	-9.6	0.011
checks bank balance most days	359	17.6	2.8	.	1.7	.	4.5	.
at least once a week	718	35.2	3.2	.	3.5	.	6.8	.
a couple of times a month	362	17.7	-0.2	.	-0.3	.	-0.5	.
at least once a month	372	18.2	-3.9	.	-3.5	.	-7.4	.
less than once a month	111	5.4	-0.6	.	0.0	.	-0.6	.
never	118	5.8	-1.4	0.000	-1.4	0.291	-2.8	0.001
checks bank balance using cashpoint/ATM (check all)	612	31.8	-1.5	0.003	-7.2	0.006	-8.7	0.002
online	965	50.2	6.7	0.000	3.3	0.196	10.0	0.000
by telephone	71	3.7	-0.1	0.553	-1.2	0.151	-1.4	0.118
app on a mobile device	403	21.0	4.8	0.000	15.1	0.000	20.0	0.000
text messages or alerts from bank	74	3.9	0.5	0.012	2.2	0.047	2.7	0.011
paper statement	415	21.6	-6.4	0.000	-5.2	0.026	-11.6	0.000
other	30	1.6	-0.5	0.004	-0.3	0.574	-0.8	0.271
did not file a tax return last year	1676	82.8	-1.2	.	2.9	.	1.7	.
filed tax return, online	248	12.3	1.6	.	-0.3	.	1.3	.
filed tax return, paper form	99	4.9	-0.4	0.000	-2.6	0.127	-3.0	0.078
no store loyalty cards	341	16.7	-1.9	0.000	-5.0	0.021	-7.0	0.004
N	2,049		1,623		267			

Notes: Difference = percentage point difference in column percentages. Difference[1] = device users minus full sample with known device status, Difference[2] = participants minus device users, Difference[3] = participants minus full sample with known device status. P-values from Chi2 tests adjusted for clustering and stratification. P-value[1] = difference between device users and non-device users, P-value[2] = difference between participants and non-participants, conditional on device use, P-value[3] = difference between participants and non-participants, full sample with known device status.

Discussion

We report on one particular implementation of research using mobile technology (a spending app to record purchases over a month) in the context of a large-scale probability household panel. Prior to inviting panel members to participate in the spending study, we measured a number of potential covariates related to access to technology, ability, and willingness to participate in the study, as well as their financial position and financial behaviours. We invited all eligible sample members to the study, regardless of their reported access to the requisite technology.

With regard to RQ1, we found that 76.3% of respondents had a mobile device. Of these, 20.2% completed the registration survey and 15.8% used the app at least once. This translated to 16.5% of all invited sample members completing the registration survey and 12.8% using the app at least once during the study. In addition, there were some people who had reported not having a mobile device but then did participate in the app study. This suggests that access to technology is fluid and should not be used as a criterion to exclude potential participants. We did not collect information on the operating system used on respondents' mobile devices, so we may have lost a few more respondents due to incompatible devices (4.3% of non-participants who had a mobile device mentioned this in the end of project survey, see Table 3).

We embedded a small incentive experiment varying the incentive to download the app (RQ2), assuming this would be the biggest barrier to participation. We did not include a "no bonus" control group because of limited power. We found no effect of the differential incentive. One interpretation is that the initial incentive was not large enough to get sample members to take the initiative to download the app. A number of non-participants reported issues relating to the process of downloading and installing the app. We speculate that this step, which requires action on the part of participants, may be a big hurdle to participation in mobile-based studies such as this. Even those studies using passive measurement require this initial step and learning more about how to overcome this initial inertia is important for studies using mobile apps.

As an aside, because of the scarcity of prior research using apps, we found it difficult to estimate the costs of the project prior to launch, and to budget an appropriate amount for the various types of incentives we used. A higher initial incentive may have increased participation, but would also have increased the costs of the project. An unconditional incentive (consistent with the literature) may have yielded more participants, but may not have been cost-effective given the relatively low participation rate. Similarly, would higher (or lower) daily incentives impact ongoing participation,

28

and would incentivising each scan rather than daily use impact the number of purchases reported? Further research on the optimal combination of incentives to maximize participation across the life of the study is needed.

Regarding RQ3, we found that, while respondents could use smartphones or tablets to download the app and scan receipts, the majority used smartphones. This is encouraging, as it allowed participants to scan receipts at the time of purchase. Among participants who have access to both devices, there is variation in hypothetical willingness and comfort using specific devices for particular tasks (see Wenz, et al. in press). Understanding these distinctions in respondent preference and use of devices is key to exploiting the benefits of mobile technologies for data collection.

RQ4 addressed reasons for non-participation among mobile device users. We found sizeable proportions of respondents reporting reasons related to the ability to use the technology, whether due to the limits of the technology itself, such as insufficient storage capacity, or to participants' confidence or ability in using the device. Understanding these barriers and finding ways to overcome them is another key challenge for research using mobile devices.

Examining the patterns of participation across the weeks of the spending study (RQ5), we saw surprisingly low drop-out out after initial use of the app. This may have been related to the bonus incentive for participating every day of the month, but also suggests that the experience of scanning was not so burdensome that it deterred people from continued participation. This interpretation is supported by findings in a companion paper by Read (2018), examining subjective and objective respondent burden in the spending study: participants spent on average less than one minute a day reporting their spending in the app, most said in the self-completion debrief questionnaire that they would be willing to participate in such a study again, and the time it took participants to use the app was not predictive of future drop-out from the study. The fact that we see little evidence of fatigue across the month of the study is in contrast to other intensive measurement studies, like expenditure and travel diaries (e.g. Schmidt 2014).

With regard to the ongoing incentive, we decided against giving a (smaller) reward for each scanned receipt or entered purchase, because we did not want to incentivise people to scan receipts that were not theirs. But the incentive for using the app at least once during the day does not seem to have incentivised people to use the app only once a day. In the registration survey, most people reported purchasing goods or services once a day or less. As noted earlier, mapping the optimal incentive onto the desired behaviour (frequency of reporting) is an area for further research.

A key contribution of our paper is the exploration of factors other than socio-demographic variables in the decision to participate in an app-based study. In RQ6 we examined various potential barriers to participation, among sample members who have mobile devices. Consistent with the elaborated view of the digital divide (see Hargittai 2002), we find that personal use of the technology for specific activities is related to participation in the app-based spending study. Frequency of mobile device use and willingness to download an app remain significant predictors in the full model controlling for a variety of other factors. However, indicators of general cooperativeness and willingness to share personal information are also significant in the full model. This suggests that both broad willingness to share data and more proximate factors related to the specific task are important in determining participation. The fact that the relationship of age (which is a strong correlate of digital access) with participation is no longer significant in the full model suggests that the more proximate ability and device use variables are more important. In contrast, the significant effect of gender (with women participating at a higher rate than men) in the full models suggests that this is not explained by gender differences in access, ability or willingness to use the technology. Potential explanations are that 1) women are generally more willing to cooperate with research requests (see chapter 5 in Groves and Couper 1998), 2) women are more likely to do the shopping, and/or 3) women are more likely to do the household budgeting or manage the finances. This is an area for future research, and suggests additional variables to measure as covariates. In addition, the finding that the behavioural and attitudinal measures remain significant predictors of participation, suggests that adjusting on socio-demographic variables alone may not be sufficient to minimize non-response bias. This parallels findings on correcting for selection bias in participation in internet surveys, where weighting based on socio-demographic variables alone may also perform poorly (see e.g. Couper, Kapteyn, Schonlau and Winter 2007, Tourangeau, Conrad and Couper 2013).

Examining the components of selection bias (RQ7), we find extensive coverage bias: mobile device users differ from non-users in socio-demographic characteristics, financial behaviours, and correlates of spending. This implies that the standard "digital divide" between haves and have-nots remains an important source of bias in mobile studies. Conditional on having a device, there is some participation bias in socio-demographic characteristics and financial behaviours. The differences are in the same direction as the coverage bias. Crucially, however, there are no differences between participants and non-participants in correlates of spending. Examining participation bias in the full sample, much of the coverage bias is washed out, since the coverage rate is so much higher than the participation rate. Overall, there are some differences in demographics between participants and non-participants. There are also differences in terms of some of the behaviours related to use of the

technology (e.g., frequency and method of checking bank balances). Those who use store loyalty cards are over-represented in the sample (see Biler, et al. 2013). However, we find no evidence of bias in terms of variables related to the outcome of interest, expenditures. Despite the relatively low participation rate, participants are no different from non-participants on several key income and spending-related indicators. This is an encouraging finding: while there is still extensive coverage bias in who has and does not have mobile devices, and while participation in the app study is low, there is no bias related (in our case) to the outcome of interest. Given contrasting findings of Armoogum et al. (2013) that participants in a GPS travel study were more frequent travellers, this is an area for further research.

In summary, our study contributes to the emerging literature on mobile technologies to enhance and extend measurement in surveys. While there is extensive coverage bias in who has and does not have mobile devices, and while participation rates in the app-based study are relatively low, most who do participate remain in the study for the full month and do not appear to be a biased sample in terms of the outcome measured by the app. Our results also suggest that as the use of mobile technologies for personal purposes increases, including among older groups in the population, participation in survey activities using these technologies is likely to increase.

References

Angrisani, M./Kapteyn, A./Samek, S. (2017). "Real time measurement of household electronic financial transactions in a population representative panel," *7th conference of the European Survey Reserach Assocation*.

Antoun, C./Conrad, F. G./Couper, M. P./West, B. T. (2018). "Simultaneous estimation of multiple sources of error in a smartphone-based survey," *Journal of Survey Statistics and Methodology*, published online first, DOI: https://doi.org/10.1093/jssam/smy002.

Archer, K. J./Lemeshow, S. (2006). "Goodness-of-fit test for a logistic regression model fitted using survey sample data," *Stata Journal*, 6(1): pp. 97-105.

Armoogum, J./Roux, S./Pham, T. H. T. (2013). "Total nonresponse of a GPS-based travel survey," *conference on New Techniques and Technologies for Statistics*.

Benzeval, M./Kumari, M./Jones, A. M. (2016). "How do biomarkers and genetics contribute to Understanding Society?," *Health Economics*, 25(10): pp. 1219-1222.

Biler, S./Šenk, P./Winklerová, L. (2013). "Willingness of individuals to participate in a travel behavior survey using GPS devices," *conference on New Techniques and Technologies for Statistics*.

Brosnan, K./Grün, B./Dolnicar, S. (2017). "Pc, phone or tablet? Use, preference and completion rates for web surveys," *International Journal of Market Research*, 59(1): pp. 35-55.

Couper, M. P./Gremel, G./Axinn, W. G./Guyer, H./Wagner, J./West, B. T. (2018). "New options for national population surveys: The implications of internet and smartphone coverage," *Social Science Research*, published online first, DOI: https://doi.org/10.1016/j.ssresearch.2018.03.008.

Couper, M. P./Kapteyn, A./Schonlau, M./Winter, J. (2007). "Noncoverage and nonresponse in an internet survey," *Social Science Research*, 36(1): pp. 131-148.

Crawford, S. D./McClain, C./Young, R. H./Nelson, T. F. (2013). "Understanding mobility: Consent and capture of geolocation data in web surveys," *68th conference of the American Association for Public Opinion Research*.

Fuchs, M./Busse, B. (2009). "The coverage bias of mobile web surveys across European countries." " *International Journal of Internet Science*, 4(1): pp. 21-33.

Groves, R. M./Couper, M. P. (1998). *Nonresponse in household interview surveys*, John Wiley & Sons.

Groves, R. M./Fowler Jr, F. J./Couper, M. P./Lepkowski, J. M./Singer, E./Tourangeau, R. (2009). *Survey methodology, 2nd edition* (Vol. 561, 2nd ed.), New York: Wiley.

Hargittai, E. (2002). "Second-level digital divide: Mapping differences in people's online skills," *First Monday*, 7(4).

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

Jäckle, A./Gaia, A./Al Baghal, T./Burton, J./Lynn, P. (2017). "Understanding Society: The UK Household Longitudinal Study Innovation Panel, waves 1-9, user manual," Colchester: University of Essex.

Keusch, F./Antoun, C./Couper, M. P./Kreuter, F./Struminskaya, B. (2017). "Willingness to participate in passive mobile data collection," *72nd conference of the American Association for Public Opinion Research*.

Lessof, C./Jäckle, A./Couper, M. P. (2017). "Data quality from a mobile app survey to collect expenditure data as part of a large-scale probability household panel survey," *7th conference of the European Survey Research Association*.

Link, M. E./Murphy, J./Schober, M. F./Buskirk, T. D./Childs, J. H./Tesfaye, C. L. (2014). "Mobile technologies for conducting, augmenting and potentially replacing surveys: Report of the AAPOR task force on emerging technologies in public opinion research," *Public Opinion Quarterly*, 78(4): pp. 779-787.

Lugtig, P./Toepoel, V./Amin, A. (2016). "Mobile-only web survey respondents," *Survey Practice*, 9(3).

Maslovskaya, O./Durrant, G./Smith, P. W. F./Hanson, T./Villar, A. (2017). "What do we know about mixed-device online surveys in the UK?," *Presented at the European Survey Research Association conference, July, Lisbon*.

Metzler, A./Fuchs, M. (2014). "Coverage error in mobile web surveys across European countries," *Presented at the Internet Survey Methodology Workshop, December, Bozen-Bolzano*.

Metzler, A./Fuchs, M. (2017). "The mobile web only population – socio-demographic characteristics and potential bias," *Presented at the European Survey Research Association conference, July, Lisbon*.

O'Doherty, K./Jaszczak, A./Hoffmann, J. N./You, H. M./Kern, D. W./Pagel, K./McPhillips, J./Schumm, L. P./Dale, W./Huang, E. S. (2014). "Survey field methods for expanded biospecimen and biomeasure collection in NSHAP wave 2," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 69(Suppl_2): pp. S27-S37.

Ofcom. (2017). "Communications market report " London: Ofcom.

Pinter, R. (2015). "Willingness of online access panel members to participate in smartphone application-based research," in *Mobile research methods: Opportunities and challenges of mobile research methodologies*, Eds. D. Toninelli, R. Pinter and P. d. Pedraza, London: Ubiquity Press, pp. 141-156.

Read, B. (2018). "Respondent burden in a mobile app: Evidence from a shopping receipt scanning study," *Understanding Society Working Paper* 2018-04. Colchester: University of Essex.

Revilla, M./Couper, M. P./Ochoa, C. (2017). "Willingness of online panelists to perform additional tasks," *General Online Research conference*.

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

Sakshaug, J. W./Couper, M. P./Ofstedal, M. B./Weir, D. R. (2012). "Linking survey and administrative records: Mechanisms of consent," *Sociological Methods & Research*, 41(4): pp. 535-569.

Scherpenzeel, A. (2017). "Mixing online panel data collection with innovative methods," in *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung*, Eds. S. Eifler and F. Faulbaum, Wiesbaden: Springer, pp. 27-49.

Schmidt, T. (2014). "Consumers' recording behaviour in payment diaries – empirical evidence from Germany," *Survey Methods: Insights from the Field*, Retrieved from http://surveyinsights.org/?p=4563.

Silberstein, A. R./Scott, S. (1991). "Expenditure diary surveys and their associated errors," in *Measurement errors in surveys*, Eds. P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz and S. Sudman, Hoboken, NJ: Wiley, pp. 303-326.

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): pp. 544-560.

Tourangeau, R./Conrad, F. G./Couper, M. P. (2013). *The science of web surveys*, New York: Oxford University Press.

University of Essex. Institute for Social and Economic Research. (2018). "Understanding Society: Spending Study 1, 2016." [data collection]. UK Data Service. SN:8348. http://doi.org/10.5255/UKDA-SN-8348.

University of Essex. Institute for Social and Economic Research. (2017). "Understanding Society: Innovation Panel, waves 1-9, 2008-2016." [data collection], UK Data Service, 8th Edition. SN: 6849, http://doi.org/10.5255/UKDA-SN-6849-9.

Wenz, A./Couper, M. P./Jäckle, A. (in press). "Willingness to use mobile technologies for data collection in a probability household panel," *Survey Research Methods*.

Appendix

Understanding Society Spending Study FAQ (from participant website)

You may have recently received a letter or email from us, asking you to take part in a new study which can help researchers look at factors that affect our income and spending, but also how our financial situation affects other parts of our lives, such as our health. Here is a list of FAQs, if your question is not here, please contact us.

When do I get my gift-card?

One week after the month is completed, we will see who has finished the study and whether they have completed the end-of-project online survey. We will process the rewards each week, and the gift-cards will be sent to you by Love2Shop within 14-21 days.

What is this study?

The Understanding Society Spending Study is a research project which is trialling new ways of collecting information on spending that are easier than detailed questions. In our last survey, we asked some new questions in order to better understand how you manage your finances. By combining this information, with the information from receipts, we will get a clearer picture of how different households manage their money.

The Understanding Society Spending Study is being conducted by researchers at the Institute for Social and Economic Research at the University of Essex, with our partners at Kantar.

I am having trouble logging in to the app, where is my user name and password?

After you complete the registration survey online, you should reach a screen which gives you your unique user-name. This will start with UK and be followed by 6 numbers, for example UK012345. You do not need a password, you can leave this blank.

Can I use my Windows phone?

Unfortunately, the app used for this study is only available for mobile devices which use the Android or the Apple iOS operating systems.

What do you want me to do?

Go to the URL printed on your letter or click the link on the email we sent you. After answering a few quick questions, you will be given information on how to download the PanelSmart app. Once you have downloaded the app and registered, when you buy something, you will be able to use the app to take a picture of your receipt and send it to us. If you spend some money and don't get a receipt, or you did not spend anything in a particular day, you can record that using the app as well. We would like you to use the app daily for a month.

Where can I get the app?

You will receive instructions after you complete the short online survey. The app is available from the Apple App Store and the Google Play app store for Android. Search for "PanelSmart". Download the app and then launch it.

How do I submit a receipt?

When you have a receipt, select the "Submit Purchase or Nothing Bought Today" option on the first menu. Then select the "Submit a Receipt" option. You will then be able to use the camera on your phone or tablet to take a picture of your receipt. There are a couple of screens with information on how to do this and then you will get a screen where you will see an icon of a camera. Press this to take a photo or to upload a photo you've already taken. You will get a chance to re-take the photo if necessary. If the receipt is long or double-sided you will be able to select an option to photograph another section of the receipt, or to indicate that the full receipt has been captured. Once that is done, you can press and the app will send us the receipt.

What about online purchases?

If you get a receipt from the online purchase (e.g., such as supermarket online shopping), you can scan that as normal. Otherwise, you can let us know about spending where you did not get a receipt using the app. To do this you should go to the "Submit Purchasing or Nothing Bought Today" option and then select the "No receipt" option.

What if I don't spend anything during a particular day?

Please tell us about this as well. To do this you should go to the "Submit Purchasing or Nothing Bought Today" option and then select the "Nothing bought" option.

How long will it take?

It only takes a few seconds to use the app, photograph a receipt and send it to us.

What's in it for me?

As a token of our appreciation for your help, we will reward you for your participation. We will keep track of a reward account, and when you download and install the app, we will add your welcome reward to the account. Each day that you use the app, even if it's to tell us you didn't spend anything that day, we will add 50p to your reward account. At the end of 31 days, if you have been active every day, you will get a bonus of £10. At the end of the survey period, if you answer a small set of questions about your experience with the app, you'll earn another £3. We will send you weekly updates to let you know how much you have earned and will send you a Love2Shop gift card for that amount at the end of the study.

What happens with my information?

We can use the information on your receipts to understand your pattern of spending over the month, as well as gathering some information about the shops you use, and the products you buy. The images you send us will be anonymised, and the items bought, and the cost, will be coded into categories, such as "food", "health and beauty", "household cleaning" and so on. The information on the name of the shop and the date and time of the purchase will also be recorded. At no point will researchers have access to any of your personal information. The information provided will only be used for research purposes.

Who should I contact if I need help?

Please email us if there are technical issues, we will pass your query on to Kantar WorldPanel and they will get back to you as soon as we can.

Which type of purchases should I submit?

Please report all money spent on buying goods and services – excluding mortgage or rent payments and regular bills (such as gas, electricity, water, council Tax, internet, telephone, mobile phone and household and car insurance). But include money spent at a point of sale (e.g., store, petrol station, restaurant, etc.), online, or for other purchases in cash, by cheque or one-off bank transfer (e.g., babysitter, workmen, vending machines, etc.).

Please include:

- Food and groceries
- Clothes and footwear
- Transport costs, e.g., petrol, car maintenance, public transport costs
- Child costs, e.g., childcare, school equipment and fees
- Home improvements and household goods, e.g., DIT, gardening, furniture, white goods or electrical goods
- Health expenses, e.g., glasses, dental care, prescriptions, social care
- Leisure and other discretionary spending:
- Socialising and hobbies, e.g., going out (restaurants, pub, cinema, theatre, concert), gym, or club membership, arts and crafts, children's activities
- Other goods and services, e.g., books, magazines, DVDs, Blu-Rays, CDs, downloads, games, toys, beauty products, haircuts, manicures, massages
- Holidays
- Giving money or gifts to other people, e.g., money for children, gifts or money for relatives, donations to charity

What if the receipt includes some items for someone else?

Please submit the receipt anyway.

What should I do if someone else in the household has a receipt but is not taking part in this Spending Study?

We will ask you to estimate the total amount of money spent by other members of your household at the end of each week. You do not need to scan their receipts.

Tips for capturing your receipts

We need to be able to read all the details on your receipts so it's very important for the pictures to be as clear as possible. If we're unable to use a receipt we may miss some important information about household spending patterns, so below are some tips to help you take the best quality pictures.

- Capture all details
- Make sure all details printed on the receipt are captured in your images, from the very top right to the bottom, but don't take the picture too far away from the receipt as the text may become too small to read.
- For particularly long receipts with approximately 30+ items, you may need to capture it in sections.
- For shorter receipts with fewer than 30 items, hold the phone at a distance where the whole length of the receipts fits just within the picture.

Creases and wrinkles

- If the receipt has been folded please try to make it as flat as possible before you take the picture. When possible we recommend taking the picture straight after your purchase, this way you also won't forget to send it.

Lighting

- Make sure there is enough light on the receipt so that the text is clear in the picture. If the light is too dim it may not be possible for us to read the text. If it is too dark, try using the camera light if your phone has one.

Perspective

- Take the picture from directly above the receipt, e.g., so that the receipt appears as a flat rectangular shape in the image and not at an angle.

Blurring

- Keep your phone held as steady as possible when you take the picture to avoid the text becoming blurred. We also suggest placing the receipt on a flat surface such as a desk or table so that it is stable.

Long receipts

- For long receipts – e.g., a grocery receipt with lots of items (30+) – please take up to four pictures, starting from the top of the receipt and working down. We suggest folding the receipt in half so you can be sure you don't miss any details in the middle. Alternatively, you can cut the receipt into parts (up to 3) and place them side-by-side to capture them all in one picture.

Double-sided receipts

When submitting your receipt, please take one image of the front of the receipt, then one of the reverse.

Multiple receipts in picture

Avoid capturing multiple receipts in the same picture, e.g., in the background or to the side of the one you are photographing.

I have submitted the same receipt twice, what should I do?

We suggest that it is best to send receipts as soon as you receive them so you don't forget. However, we can identify duplicate receipts and remove them from our data so you don't have to do anything if you make a mistake.

Can I check which receipts I've already sent?

Unfortunately, it is not currently possible to see details of which receipts you have sent us. We hope to provide an option to view this in the future.

Do I need to send receipts as soon as I make a purchase?

It is not necessary to send immediately after a purchase but please try to send as soon after your trip as possible so you do not forget.

My receipt shows credit card details, how can I remove them?

You can cover the card number or blank it out using a pen, but please do not cover any details about the items purchase, price or the store or date. Also, please do not cut off the bottom of the receipt as often this includes the date and time of the trip which is very important to us.

How much data does it take to send a receipt?

The amount of data required to send an image depends on your phone's camera resolution. Most smartphone cameras typically have a resolution of 3 megapixels or higher. Images at 3 megapixels will be around 500 kilobytes (0.5 megabytes). On some smartphones it is possible to adjust the resolution of the camera, if so please set the camera to use at least 3 megapixels.

Can I transmit by Wi-Fi only?

Yes. If you have a low data allowance on your mobile phone plan you can choose to transmit data over Wi-Fi only. In the PanelSmart app home screen press the 'Menu' button then select the 'Settings' option and check the box for "Wi-Fi only". If you choose this option please remember to connect to a wireless network regularly in order to send your data.