Understanding Society

Working Paper Series

No. 2017 – 09

October 2017

Participation in a Mobile App survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases

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

Understanding Society
THE UK HOUSEHOLD LONGITUDINAL STUDY

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

Non-technical Summary

As part of a broader project aimed at looking at ways to improve the measurement of financial information in *Understanding Society*, we are exploring the role that new technologies may play in making it easier for survey participants to provide information on expenditures. We used a spending app developed by Kantar Worldpanel that captures images of shopping receipts and allows participants to record other types of expenditures. In the first test of the acceptance and use of this kind of app in a probability survey, we invited over 2,400 *Understanding Society* Innovation Panel wave 9 (IP9) respondents to download the app and scan receipts and report expenditures over a month. We invited all eligible IP9 respondents, regardless of whether or not they had a suitable mobile device.

This paper focuses on how many participated in various parts of the task, and how they differed from those who did not participate. Among IP9 respondents, 16.5% installed the app and completed the short registration survey, while 12.8% used the app at least once. Most of those who participated used a smartphone (86.5%), while 11.5% used a tablet and the balance used both. Major reasons mentioned for not participating included a variety of technical difficulties (including not having a suitable device), time constraints, lack of confidence using mobile devices or apps, and concerns about sharing detailed expenditure information. Among those who used the app at least once, most continued to do so over the course of the month, with 81.5% remaining in the study for at last 28 days.

Using data from the IP9 interview, we also explored differences between participants and non-participants. We found that those who use their mobile devices more frequently were more likely to participate in the spending app study, as were those who were generally more cooperative in the survey. Those who use technology to manage their finances and who use store loyalty cards were also more likely to participate. However, we found the two groups to be similar on several income and spending-related indicators.

Participation in a Mobile App survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases

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

Abstract:

We examine non-response in expenditure data collected with a mobile app: 2,432 members of the *Understanding Society* Innovation Panel in the UK were invited to download an app to record all their spending on goods and services for a month, by scanning receipts or reporting spending in the app. We examine participation at different stages of the process; the prevalence of potential barriers to participation, including access, ability and willingness to use different mobile technologies; and biases in the types of people who participate, considering socio-demographic characteristics, financial position and financial behaviours.

Keywords: mobile app, scanning, finances

JES Classification: D14, C81, C83

Acknowledgements: This research was funded by the UK Economic and Social Research Council (ESRC) and the National Centre for Research Methods (NCRM) (ES/N006534/1). Carli Lessof received funding from the ESRC Doctoral Training Programme. The *Understanding Society* Innovation Panel is funded by the ESRC (ES/N00812X/1). We are grateful for the in-kind contributions of our project partner Kantar Worldpanel who implemented the spending study, to Brendan Read for excellent research assistance, and the NCRM International Visitor Exchange Scheme for funding research visits by Mick Couper to the University of Essex.

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 technologies, 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. Similarly, 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 et al. 2017b). 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 both the rate of non-response at various stages and the correlates of such non-response.

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. Benzeval, Kumari and Jones 2016; O'Doherty et al. 2014; Sakshaug et al. 2012). 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 my further increase the opportunities for non-participation.

Mobile technologies bring a range of additional tasks that can be performed. These vary in the degree of respondent involvement (from installing an app that passively collects and transmits data to actively using a device to provide repeated measurements), the level of burden (from one-time measurement to frequent measurements over time), the potential threats to privacy (i.e., the sensitivity of the data being collected), and the technical requirements of the task (e.g., battery usage, volume of data transmission) (see Wenz, Couper and Jäckle 2017). As the complexity of the task expands and as the number of measurement opportunities increases, the issue of non-response becomes more complex.

There are a number of different barriers that may affect participation in an app-based activity and potentially 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, 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 and motivation to participate in surveys, 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 this 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 existing literature is still very sparse and few studies have examined the rates of participation at each stage and the nature of selection biases that may result.

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 UK Innovation Panel, Wenz, Couper and Jäckle (2017) 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. A few papers have explored actual compliance with 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).

A few papers have explored factors related to non-response or examined potential non-response bias (e.g. Armoogum, Roux and Pham 2013; Biler, Šenk and Winklerová 2013; Keusch et al. 2017; Pinter 2015; Revilla, Couper and Ochoa 2017; Revilla et al. 2016; Wenz, Couper and Jäckle 2017). 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. (2017) 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 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 non-response, from 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 proportion of a general population sample participates in the app study? Do incentives increase participation? Do survey non-respondents engage in the app study?
2. Which devices do participants use and does device choice correspond to previously stated preferences?
3. What are the main reasons for not participating in the app study?
4. What are the patterns of participation over the month?
5. How prevalent are potential barriers to participating in the app study? Which are most important in predicting participation?
6. What is the nature of non-participation bias? Is non-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 et al. 2017a 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. 2017a) and the number of full individual interviews was 2,114. 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.[1] 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. 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 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

[1] The data and documentation from the spending study will be available from the UK Data Service.

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. 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 (iTunes 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 sent an end of project questionnaire, with questions tailored to ongoing participants, participants who dropped out before the end of the month, and non-participants. Non-respondents to the email survey were sent a questionnaire by post, with a Freepost return envelope, but no incentive. The response rate for the end of project survey was 70.4% for those who used the app at least once and 32.4% for non-participants.

Outcomes: measures of participation

The outcome measures are derived from the app paradata which recorded a total of 11,561 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.[2] 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. Respondents who did not complete the self-completion module are dropped from the analysis of predictors of participation. These include face-to-face respondents who declined to do the self-completion section (N=44) and telephone respondents who by design were not asked the self-completion questions (N=29). For the analysis of predictors of participation, the resulting analysis sample includes 2,041 IP9 respondents, of which 268 participated in the spending study.

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.
- ***Has a smartphone, has a tablet:*** derived from a 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,

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

and 2 if not, or if the answer is missing by design because the respondent does not use the internet for personal use.

- ***Wifi at home:*** coded as 1 if respondent has wifi access at home, and 2 if not, or if the answer is missing by design because the respondent does not use the internet.
- ***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 do not have either, or the answer is missing by design because they do not use the internet or do not have a smartphone.

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 respondent uses at least one of the devices every day, and 2 if less often or not asked because they do not have a smartphone or tablet.
- ***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 or not asked because they do not have a mobile device.
- ***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 or if they do not have a mobile device.

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, or not asked because they do not have a mobile device.

- ***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, of if the question was not asked because the respondent does not have a mobile device.

Additional variables related to *general cooperativeness* with the survey 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 item non-response rate for predictors of participation due to "don't know" and "refused" responses were mostly <0.5% (19 items), below 2.0% for a further 10 items, and 2.5% for the consent to linkage question. Due to the low rates of missingness we treat these respondents in the same way as respondents for whom responses are missing by design due to the routing of the questionnaire.

All variables used for the analysis of non-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 personal income, household spending on food, and household spending on fuel we treat missing observations as a separate category (see Table 5). For all other items the 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 proportion of a general population sample participates in the app study? Do incentives increase participation? Do survey non-respondents engage in the app study?

Among IP9 respondents 16.5% completed the registration survey and 12.8% used the app at least once. Subsequent drop-out was unexpectedly low: 10.2% used the app at least once in each of the five consecutive weeks (Table 1). The incentive experiment, varying the value of the bonus for downloading the app, had no effect on these participation outcomes. Not everyone in the sample has a mobile device though: among the 76.2% of the sample who do have a smartphone or tablet, 20.2% completed the registration survey and 15.8% used it at least once.

Table 1: Participation in the spending study

	N	%	£2 bonus (%)	£6 bonus (%)	P-value
Issued sample (IP9 respondents)	2,114	100.0	--	--	--
Completed the registration survey	348	16.5	15.9	16.9	0.595
Used app at least once	270	12.8	11.9	13.6	0.326
Used app at least once in each of five weeks	216	10.2	9.4	10.9	0.237

Notes: P-values from Chi2 tests adjusted for clustering and stratification.

We also invited IP9 non-respondents, living in households with at least one IP9 respondent, to participate in the app study (N=318). 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 1.3% 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.

RQ2: Which devices do participants use and does device choice correspond to previously stated preferences?

The majority of participants 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

RQ3: What are the main reasons for not participating in the app study?

Table 3 shows the responses given by non-participants 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. Among the 598 non-participants who answered the end of project survey, the most frequent single response was that respondents did not have time to scan (40.4%). About two-thirds (64.8%) of responses 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 (42.4%) mentioned lack of confidence, either with using their mobile device for this kind of activity and/or with downloading apps. Finally, 40.7% mentioned at least one privacy concern: they were not willing to share spending information, or not confident that information would be held securely, and 13.4% said they were not interested.

Table 3: Reasons for not participating in the app study

	N	% of cases
Did not have time to scan	199	40.4
Did not try to download the app	160	32.5
Do not have a smartphone or tablet which can download apps	140	28.4
Not confident using my phone or tablet for this kind of activity	108	21.9
Not willing to share spending information	107	21.7
Not able or confident to download apps onto my phone or tablet	105	21.3
Not confident that information would be held securely	81	16.4
Do not have access to the internet on my phone or tablet	68	13.8
Not interested	66	13.4
Did not have sufficient storage space to download the app	59	12.0
Could not download the app because not compatible with operating system	34	6.9
Could not find the app in the app store	23	4.7
Link to downloading the app did not work	22	4.5

Notes: N=598. Multiple mentions.

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. 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.

[3] Receipts where the date on the receipt preceded the day on which the invitation to the spending study was sent out (N=43) were dropped from Figure 1 and Figure 2, although results are unchanged if they are included.

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

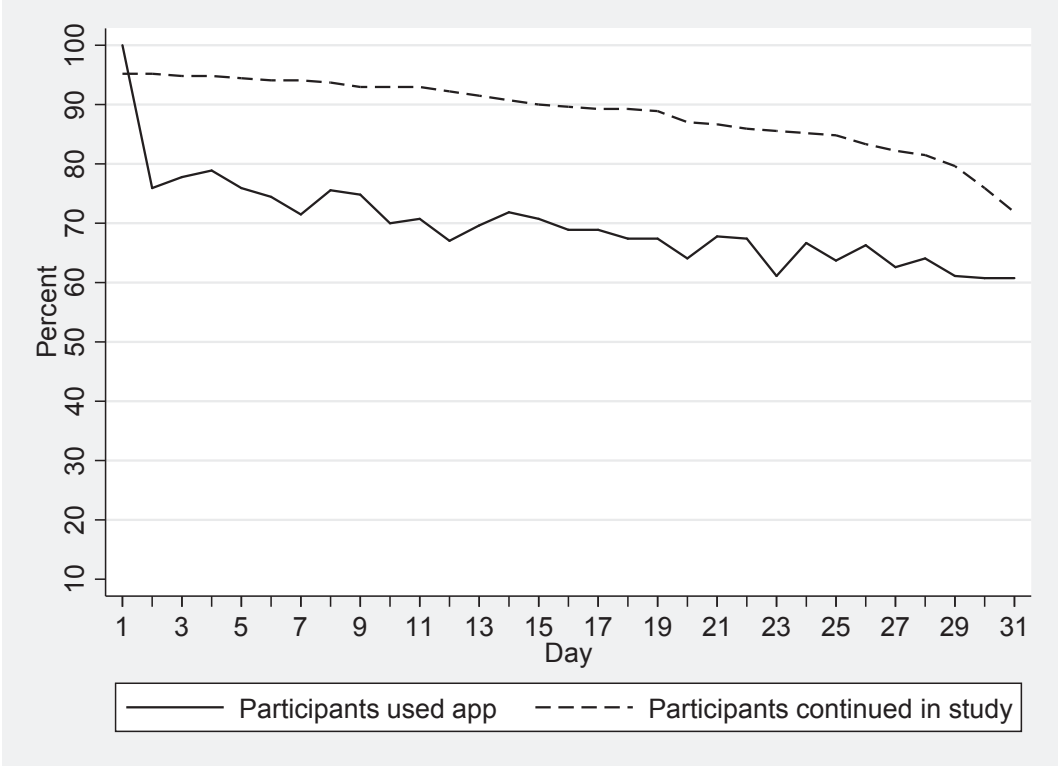
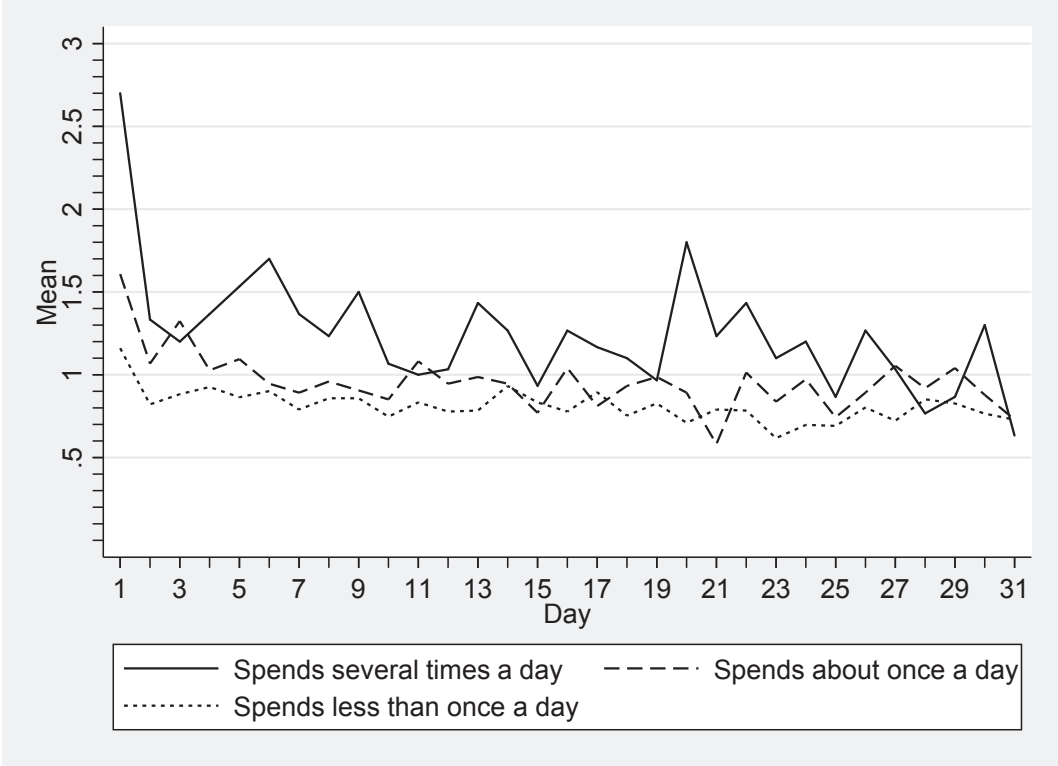


Figure 2: Mean number of receipt scans and purchases entered, by self-reported frequency of spending money on goods and services



RQ5: How prevalent are potential barriers to participating in the app study? Which are most important in predicting participation?

Table 4 shows the prevalence of different potential barriers relating to access, ability and hypothetical willingness to use the spending study app, among all IP9 respondents. Column 1 shows that barriers to *access* remain high: although 73.1% use the internet daily, only 65.4% have a smartphone and 60.2% have a tablet. Barriers related to *ability* to participate in an app study are similarly high: 65.9% report that they use at least one of their devices daily, 46.3% consider themselves advanced users, and between 46.6% and 69.8% use at least one of their devices to take photos, make online purchases, use online banking or install apps. Barriers related to *willingness* are also high: only 40.6% say they would be very or somewhat willing to download an app and 49.1% to use the camera on either device for a survey. Only 16.4% would not at all be concerned about the security of providing information by downloading an app to complete an online questionnaire and 27.1% 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, for which item non-response and consent to data linkage from the IP9 interview are used as indicators. The item non-response rate ranged from 1.8% to 31.1%, with a median of 4.0%. Consent to data linkage was given by 57.5% of the sample. As consent was lower among respondents who completed their questionnaire online, we control for the mode of interview in the regression models: 54.0% of respondents completed their interview online, 46.0% completed in face-to-face interviews.

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 significance level of the Chi2 test is $P < 0.001$. 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 once 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 Pr>F ranging from 0.867 to 0.998 for each of the four partial models and the full model.

Of the predictors related to *access*, using the internet every day increases the probability of participating by 10.1 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 = 0.003 (Column 4). Having a smartphone and having a tablet increase the probability by 7.8 and 5.1 percentage points respectively. Of the predictors relating to *ability*, using the device every day increases the probability of participating by 5.3 percentage points, using at least one device for online banking increases it by 4.1 percentage points, and installing apps by 5.0 percentage points. Self-rated skill is also a significant predictor according to the joint F-test of whether both coefficients equal zero (Prob>F = 0.025). Of the predictors related to *willingness*, being very or somewhat willing to download an app for a survey increase the probability of participation by 8.0 percentage points and having no or little concerns about the security of providing data via an app increases the probability of participation (Prob>F = 0.029). Finally, the indictors of *general cooperativeness* with the survey each increase the probability of participating by between 6.1 and 7.6 percentage points.

In the full model (Column 5) the only predictors that remain significant are using one of the devices daily (+5.4 percentage points) and being very or somewhat wiling to download an app for a survey (+4.5 percentage points). Self-reported skill also remains significant (Prob>F = 0.032), however after controlling for frequency of internet and device use and for the ability variables related to actual use, the sign of the coefficients is reversed. 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 (+2.6 percentage points, P=0.026), the probability of participating decreases monotonically with age (Prob>F < 0.001), 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	73.1	16.4	0.101***		0.065	
	several times a week	11.8	7.1	0.036		0.042	
	several times a month or less	15.2	1.9	.	0.003	.	0.291
Has a smartphone	yes	65.4	17.3	0.078*		0.034	
	no or not asked	34.6	5.2	.		.	
Has a tablet	yes	60.2	16.9	0.051**		0.028	
	no / not asked	39.8	7.5	.		.	
Wifi at home	yes	86.5	14.7	0.019		-0.002	
	no or not asked	13.5	2.9	.		.	
Data plan	fixed data plan	55.0	17.1	-0.011		-0.036	
	pay-as-you-go contract	7.0	18.9	0.017		0.014	
	neither or not asked	38.0	6.3	.	0.523	.	0.144
Ability							
Frequency of device use	every day	65.9	17.8	0.053*		0.054*	
	less often or not asked	34.1	4.2	.		.	
Self-reported skill	advanced	46.3	20.2	0.050		-0.038	
	medium	29.3	10.2	0.003		-0.069	
	beginner or n/a	24.4	3.2	.	0.025	.	0.032
Takes photos	yes	69.8	17.0	0.025		0.016	
	no or missing	30.2	4.2	.		.	
Online purchases	yes	55.1	18.3	-0.002		-0.018	
	no or missing	44.9	6.8	.		.	
Online banking	yes	46.6	20.0	0.041*		0.029	
	no or missing	53.4	7.2	.		.	

Installs apps	yes	56.3	19.1	0.050**		0.036	
	no or missing	43.7	5.4	.		.	
Willingness							
Willingness to download app	very or somewhat willing	40.6	21.5	0.080***		0.045*	
	a little or not willing	59.4	7.4	.		.	
Willingness to use camera	very or somewhat willing	49.1	18.3	0.005		-0.021	
	a little or not willing	50.9	8.1	.		.	
Security concerns:	not at all concerned	16.4	23.1	0.073**		0.044	
complete online via app	a little or somewhat concerned	42.3	16.6	0.045		0.029	
	very or extremely concerned	41.4	5.7	.	0.029	.	0.215
Security concerns:	not at all concerned	27.1	19.9	0.050*		0.008	
use camera for barcodes	a little or somewhat concerned	36.3	15.8	0.042*		0.003	
	very or extremely concerned	36.6	5.5	.	0.064	.	0.937
General cooperativeness							
Item non-response rate > median	low item non-response	50.0	17.1	0.061***		0.039*	
	high item non-response	50.0	9.1	.		.	
Consent to data linkage	yes	57.5	15.9	0.070***		0.056***	
	no or don't know or refused	42.5	9.3	.		.	
Mode of IP9 interview	face-to-face	46.0	9.6	.		.	
	web	54.0	16.2	0.076***		0.061***	

Notes: N=2,041. All P-values from Chi2 tests corresponding to column 2 are $P < 0.001$. AME = average marginal effects estimated from probit models of probability of using app at least once. Standard errors adjusted for clustering and stratification. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

RQ6: What is the nature of non-participation bias? Is non-participation related to financial behaviours and outcomes?

Table 5 shows the difference in characteristics of the full sample of IP9 respondents and spending study participants who used the app at least once. The P-values are from Chi2 tests of the difference between participants and non-participants, adjusted for clustering and stratification. As indicated by the analysis of barriers to participation, women are over-represented in the participant sample (+5.5 percentage points), as are younger age groups (+6.0 for those age 18-30, +9.4 for those age 31-40) and those with higher education (+6.9 for those with a degree).

The app is intended to measure the value of purchases of goods and services over a month. We therefore first examine biases in measures of income and spending from the IP9 interview that should be highly correlated with the outcomes captured by the app. We see no difference between spending study participants and non-participants in their personal monthly income, the amount the household spent on food purchases in the last month, the household annual expenditure on fuel, whether the household is struggling or behind with paying housing costs or utilities, or the individual's subjective assessment of how well they are getting by financially. There are however large biases in the participant sample related to financial behaviours: people who do not keep a budget are under-represented by 10.1 percentage points, while those who keep a budget on a computer document or spreadsheet are over-represented by 6.5 percentage points. People who check their bank balance at least once a week are over-represented in the spending study. Strikingly, although not surprisingly, people who check their bank balance using an app on a mobile device are over-represented by 20.2 percentage points, people who check online by 10.1 percentage points, while those who check using paper statements or cashpoints are under-represented by 11.7 and 8.6 percentage points. There are no observed biases as to whether people filed a tax return or got help with their finances from a family member or friend (not shown), or whether they have an accountant (not shown). However people who do not have store loyalty cards are under-represented by 7.2 percentage points.

Table 5: Non-participation bias

		Full sample		Participants			
		N	%	N	%	Difference	P-value
Gender	male	954	45.1	107	39.6	-5.5	.
	female	1160	54.9	163	60.4	5.5	0.011
Age	16-30	343	16.2	60	22.2	6.0	.
	31-40	279	13.2	61	22.6	9.4	.
	41-50	368	17.4	60	22.2	4.8	.
	51-60	437	20.7	50	18.5	-2.2	.
	61-70	375	17.7	25	9.3	-8.5	.
	71+	312	14.8	14	5.2	-9.6	0.000
Highest qualification	degree	528	25.0	86	31.9	6.9	.
	GCSE, A-level, other higher	1267	59.9	161	59.6	-0.3	.
	other, none or missing	319	15.1	23	8.5	-6.6	0.001
Personal monthly income	quartile 1	452	21.4	52	19.3	-2.1	.
	quartile 2	452	21.4	52	19.3	-2.1	.
	quartile 3	452	21.4	69	25.6	4.2	.
	quartile 4	452	21.4	64	23.7	2.3	.
	missing	306	14.5	33	12.2	-2.3	0.262
HH monthly spend on food	quartile 1	509	24.1	54	20.0	-4.1	.
	quartile 2	476	22.5	63	23.3	0.8	.
	quartile 3	494	23.4	64	23.7	0.3	.
	quartile 4	488	23.1	76	28.1	5.1	.
	missing	147	7.0	13	4.8	-2.1	0.180
HH monthly spend on fuel	quartile 1	472	22.3	74	27.4	5.1	.
	quartile 2	598	28.3	74	27.4	-0.9	.
	quartile 3	343	16.2	44	16.3	0.1	.
	quartile 4	471	22.3	44	16.3	-6.0	.
	missing	230	10.9	34	12.6	1.7	0.155

Behind or struggling to pay housing costs or bills	yes	256	12.2	30	11.2	-1.1	.
	no	1834	87.8	238	88.8	1.1	0.660
Subjective financial situation	living comfortably	644	30.7	77	28.5	-2.2	.
	doing alright	865	41.2	120	44.4	3.2	.
	just getting by or finding it difficult	589	28.1	73	27.0	-1.0	0.585
How do you keep your budget?	on paper	538	25.8	76	28.3	2.4	0.375
(check all that apply)	on computer document or spreadsheet	307	14.7	57	21.2	6.5	0.007
	personal budget software on computer/laptop	22	1.1	8	3.0	1.9	0.002
	online budget programme	6	0.3	1	0.4	0.1	0.785
	personal budget app	29	1.4	7	2.6	1.2	0.103
	do not keep a budget	1233	59.2	132	49.1	-10.1	0.008
How often do you check your bank balance?	most days	372	17.7	60	22.2	4.5	.
	at least once a week	739	35.2	114	42.2	7.0	.
	a couple of times a month	376	17.9	46	17.0	-0.9	.
	at least once a month	380	18.1	29	10.7	-7.4	.
	less than once a month	114	5.4	13	4.8	-0.6	.
	never	119	5.7	8	3.0	-2.7	0.001
How do you check your bank balance?	cashpoint/ATM	631	31.9	61	23.3	-8.6	0.002
(check all that apply)	online	993	50.2	158	60.3	10.1	0.000
	by telephone	74	3.7	6	2.3	-1.4	0.099
	app on a mobile device	409	20.7	107	40.8	20.2	0.000
	text messages or alerts from bank	75	3.8	17	6.5	2.7	0.011
	paper statement	427	21.6	26	9.9	-11.7	0.000
	other	31	1.6	2	0.8	-0.8	0.261
Did you file a tax return or self-assessment	no	1724	82.8	227	84.7	1.9	.
form last year?	yes, online	257	12.3	36	13.4	1.1	.
	yes, paper form	101	4.9	5	1.9	-3.0	0.080
Which story loyalty cards do you have?	no store loyalty cards	354	16.9	26	9.7	-7.2	0.003

Notes: P-values from Chi2 tests of difference between participants and non-participants, adjusted for clustering and stratification.

Discussion

We report on one particular implementation of research using mobile technology (a spending app to report 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 16.5% of those invited completed the registration survey and 12.8% used the app at least once during the study. Rates of participation were higher for those who reported having a suitable mobile device: 20.2% and 15.8% respectively. But even some who reported in IP9 not having a mobile device participated in the app study (4.6%). 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 (7.4% of non-participants mentioned this in the end of project survey).

We embedded a small incentive experiment varying the incentive to download the app, 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 conclusion 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 of this kind.

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 also 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, 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 RQ2, 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 (see Lessof, Jäckle and Couper 2017). Among participants who have access to both devices, there is variation in hypothetical willingness and comfort using specific devices for particular tasks (see Wenz, Couper and Jäckle 2017). Understanding these distinctions in respondent preference and use of devices is a key to exploiting the benefits of the mobile technologies for data collection.

RQ3 addressed reasons for non-participation. 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 (e.g., do not have a mobile device, 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 (RQ4), 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 suggests that the experience of scanning was not so burdensome that it deterred people from continued participation. Conditional on the relatively low initial response, we see little evidence of fatigue across the month of the study, 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 RQ5 we examined the bivariate and multivariate associations of various potential barriers with participation. 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 internet use, 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 are also significant in the full model. This suggests that both broad willingness to participate in research 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 models suggest 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 et al. 2007; Tourangeau, Conrad and Couper 2013).

Finally, in terms of potential for non-response bias (RQ6), we find some differences in demographics between participants and non-participants in the spending study. We also find 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, Šenk and Winklerová 2013). However, we find very little evidence of bias in terms of variables related to the outcome of interest, expenditures. Despite the relatively low participation rate, those who participated in the study do not differ from those who don't on several key income and spending-related indicators. This is an encouraging finding suggesting that while participation in an app-based study such as this is related to ability and willingness to use mobile devices, it is not related (in our case) to the outcome of interest. Given the finding 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 the proportion of a general population sample who participate in the app-based study is 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

Archer, Kellie J, and Stanley Lemeshow. 2006. "Goodness-of-Fit Test for a Logistic Regression Model Fitted using Survey Sample Data." *Stata Journal* 6(1):97-105.

Armoogum, J., S. Roux, and T.H.T. Pham. 2013. "Total Nonresponse of a GPS-based Travel Surveys." in *Paper presented at the conference on New Techniques and Technologies for Statistics*. Brussels.

Benzeval, Michaela, Meena Kumari, and Andrew M Jones. 2016. "How Do Biomarkers and Genetics Contribute to Understanding Society?" *Health economics* 25(10):1219-22.

Biler, S., P. Šenk, and L. Winklerová. 2013. "Willingness of Individuals to Participate in a Travel Behavior Survey using GPS Devices." in *Paper presented at the conference on New Techniques and Technologies for Statistics*. Brussels.

Couper, M.P., A. Kapteyn, M. Schonlau, and J. Winter. 2007. "Noncoverage and Nonresponse in an Internet Survey." *Social Science Research* 36(1):131-48.

Crawford, S.D., C. McClain, R.H. Young, and T.F. Nelson. 2013. "Understanding Mobility: Consent and Capture of Geolocation Data in Web Surveys." in *Paper presented at the annual meeting of the American Association for Public Opinion Research*. Boston.

Groves, Robert M, and Mick P Couper. 1998. *Nonresponse in Household Interview Surveys*: John Wiley & Sons.

Hargittai, Eszter. 2002. "Second-Level Digital Divide: Mapping Differences in People's Online Skills." *First Monday* 7(4).

Jäckle, Annette, Alessandra Gaia, Tarek Al Baghal, Jonathan Burton, and Peter Lynn. 2017a. "Understanding Society: The UK Household Longitudinal Study Innovation Panel, Waves 1-9, User Manual ". Colchester: University of Essex.

Jäckle, Annette, Alessandra Gaia, Carli Lessof, and Mick P. Couper. 2017b. "Improving Survey Measurement of Household Finances: A Review of New Technologies and Data Sources." Colchester: University of Essex. Unpublished manuscript.

Keusch, Florian, Christopher Antoun, Mick P. Couper, FraukeKreuter, and Bella Struminskaya. 2017. "Willingness to Participate in Passive Mobile Data Collection." in *72nd annual conference of the American Association for Public Opinion Research*. New Orleans, LA.

Lessof, Carli, Annette Jäckle, and Mick P. Couper. 2017. "Data Quality from a Mobile App Survey to Collect Expenditure Data as Part of a Large-Scale Probability Household Panel Survey." in *Presented at the 7th European Survey Research Association Conference*. Lisbon.

Link, Michael E., Joe Murphy, Michael F. Schober, Trent D. Buskirk, Jennifer Hunter Childs, and Casey Langer Tesfaye. 2014. "Mobile Technologies for Conducting, Augmenting and Potentially Replacing Surveys: Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research." American Association for Public Opinion Research.

O'Doherty, Katie, Angela Jaszczak, Joscelyn N Hoffmann, Hannah M You, David W Kern, Kristina Pagel, Jane McPhillips, L Philip Schumm, William Dale, and Elbert S Huang. 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):S27-S37.

Pinter, R. . 2015. "Willingness of Online Access Panel Members to Participate in Smartphone Application-Based Research." Pp. 141-56 in *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, edited by D. Toninelli, R. Pinter, and P. de Pedraza. London: Ubiquity Press.

Revilla, M., M.P. Couper, and C. Ochoa. 2017. "Willingness of Online Panelists to Perform Additional Tasks." in *Paper presented at the General Online Research conference*. Berlin.

Revilla, M., D. Toninelli, C. Ochoa, and G. Loewe. 2016. "Do Online Access Panels Need to Allow and Adapt Surveys to Mobile Devices? ." *Internet Research* 26(5):1209–27.

Sakshaug, Joseph W, Mick P Couper, Mary Beth Ofstedal, and David R Weir. 2012. "Linking Survey and Administrative Records: Mechanisms of Consent." *Sociological Methods & Research* 41(4):535-69.

Scherpenzeel, Annette. 2017. "Mixing Online Panel Data Collection with Innovative Methods." Pp. 27-49 in *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung*, edited by S. Eifler and F. Faulbaum. Wiesbaden: Springer.

Schmidt, Tobias 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, Adriana R., and Stuart Scott. 1991. "Expenditure Diary Surveys and their Associated Errors." Pp. 303-26 in *Measurement errors in surveys*, edited by P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman. Hoboken, NJ: Wiley.

Toepoel, Vera, and Peter Lugtig. 2014. "What Happens if You Offer a Mobile Option to Your Web Panel? Evidence from a Probability-Based Panel of Internet Users." *Social Science Computer Review* 32(4):544-60.

Tourangeau, R., F.G. Conrad, and M.P. Couper. 2013. *The Science of Web Surveys*. New York: Oxford University Press.

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, Alexander, Mick P. Couper, and Annette Jäckle. 2017. "Willingness to Use Mobile Technologies for Data Collection in a Probability Household Panel." Colchester: University of Essex. Unplublished manuscript.

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 “WiFi only”. If you choose this option please remember to connect to a wireless network regularly in order to send your data.