

**Understanding Society
Working Paper Series**

No. 2019 – 04

June 2019

**Increasing participation in a mobile app study: the effects
of a sequential mixed-mode design and in-interview
invitation**

Annette Jäckle (University of Essex)

Alexander Wenz (University of Mannheim)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Michigan)

Non-technical summary

In this paper we examine two potential ways of increasing participation in app-based data collection: 1) inviting respondents to the mobile app study within an interview rather than by post, and 2) offering a browser-based alternative to the mobile app. These features were developed to address barriers to participation identified in a previous mobile app-based study, where we found that participation was limited by sample members not having devices that were compatible with the app, not being willing to download an app, or not feeling confident about doing such a task. The present study (Spending Study 2) required participants to download an app or use a browser-based alternative and use those to report all their purchases for one month, by entering amounts and categories of spending.

The data for Spending Study 2 were collected in May to December 2018 using two samples: the *Understanding Society* Innovation Panel and an online access panel managed by Lightspeed UK. The Innovation Panel (IP) is a probability sample of households in Great Britain with annual interviews of all adult household members. All respondents to the 2018 (wave 11) interview were invited to the app study, yielding an analysis sample of 2,638 IP respondents of whom 524 participated in Spending Study 2. In the access panel 2,878 respondents completed the baseline questionnaire, of which 1,065 participated in Spending Study 2.

The results suggest that if respondents complete the IP interview with an interviewer they are more than 3 times as likely to participate in the app study if they are invited to download the app while the interviewer is present than if they are sent an invitation by post. In contrast, if respondents complete the IP interview as a self-completion survey online, it makes no difference whether they are invited to the app study within the web survey or later by post. Using data collected in the preceding Innovation Panel interview to compare participants and non-participants in terms of socio-demographics, mobile device access and usage, and financial behaviours shows that although the mode of invitation to Spending Study 2 affected participation rates, it did not alter the composition of participant samples.

Offering a browser-based alternative can be effective: in the access panel where the browser-based alternative was offered immediately once the app had been declined, it increased study participation from 14% who used the app at least once to 37% who used either the app or the browser-based version at least once. Using data collected in the baseline questionnaire to compare participants and non-participants in terms of socio-demographic characteristics, mobile device access and usage, and financial behaviours shows that the browser-based follow-up reduced biases in the composition of participant samples. In contrast, in the Innovation Panel, where respondents were sent several reminders to download the app before being invited to the browser-based alternative, 16.9% used the app to report at least one purchase and only 3% used the browser-based version. The difference between the two designs highlights trade-offs between maximising uptake of the app and uptake of the follow-up mode.

The take home message is that the success of using mobile apps for data collection hinges on the protocols used to implement the app, such as protocols for invitations and protocols for following up non-participants.

Increasing participation in a mobile app study: the effects of a sequential mixed-mode design and in-interview invitation

Annette Jäckle (University of Essex)

Alexander Wenz (University of Mannheim)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Michigan)

Abstract

In this paper we examine two potential ways of increasing participation in app-based data collection: 1) inviting respondents to the mobile app study within an interview rather than by post, and 2) offering a browser-based alternative to the mobile app. We use experimental data from Spending Study 2, collected on the *Understanding Society* Innovation Panel and on an online access panel managed by Lightspeed UK. The results suggest that inviting respondents to an app study within a face-to-face interview increases participation, but does not bring in different types of participants. In contrast the browser-based alternative can both increase participation and reduce biases in who participates.

Keywords: spending diary, smartphone, mobile application

JEL classification: C81, C83

Acknowledgements: This research was funded by grants from the UK Economic and Social Research Council (ESRC) and the National Centre for Research Methods (NCRM) (ES/N006534/1) and by ESRC funding for the *Understanding Society* survey (ES/N00812X/1). We thank Thomas F. Crossley for comments on an earlier version of this paper.

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

1 Introduction

Mobile apps are an attractive and versatile method of collecting data. Since most people carry their smartphones with them throughout the day, mobile apps can be used to collect data passively, for example about the person's geolocation or movements; to collect data in real time, for example by triggering surveys via geo-fences or at randomized times during the day; and to implement diaries that are ideally completed close in time to the events of interest. However in samples of the general population, participation in app-based data collection is still rather low. This is in part due to coverage issues: not everyone will have a mobile device that is compatible with the study app; in part it is due to the reluctance of sample members to download and use apps for data collection; and in part this is likely because we have not yet figured out how best to implement app-based data collection in ways that reduce the barriers to participation. In this paper we examine two aspects of how an app study can be implemented, focusing on the effects on participation rates and biases.

This study is based on learning from an earlier study, in which we invited sample members to download an app on their smartphone or tablet, and to use it for one month to record all purchases of goods and services. Participants were asked to take photos of their shopping receipts, which were automatically uploaded. The idea was that this would be an easier task than manually entering information about purchases in the app. The app did, however, allow manual entry of purchases, to accommodate purchases for which participants might not have receipts. This study, referred to as Spending Study 1, was implemented in the *Understanding Society* Innovation Panel, a probability sample of households in Great Britain (University of Essex. Institute for Social and Economic Research 2019b). Sample members were sent a letter inviting them to the app study in the autumn of 2016, around three months after the end of the previous wave of panel interviews.

The findings from this first study were generally positive, but also highlighted key challenges in implementing data collection with mobile apps. The participation rate was low: only 13% of the full sample used the app, which corresponded to 16% of mobile device owners. However those who did participate were very committed and reported low burden: 82% of participants remained in the study for more than 29 days and 67% said that they would participate in a similar study again (Jäckle et al. 2019b; Read 2019). It took participants slightly more time to photograph receipts than to enter purchase information manually (on average 41 versus 30 seconds per entry) and over time participants became slightly more likely to enter purchases manually than by photographing receipts (Lessof et al. 2019; Read 2019). There were clear biases in the types of people who did and did not participate in the app-based study, which compounded differences between the types of people who did and did not have a mobile device (Jäckle et al. 2019b). There were differences in socio-

demographic characteristics, however these were mainly driven by mobile device access and usage patterns. Whether or not a sample member already did similar types of activities for their own purposes was highly predictive of whether or not they participated in Spending Study 1. For example, those who kept a budget were over-represented among participants by about 10 percentage points compared to the full sample, those who used an app on a mobile device to check their bank balance were over-represented by 20 percentage points. There were however no differences between participants and non-participants in correlates of what the app was designed to measure: household spending, whether they were struggling or behind with paying their housing costs, and subjective assessment of how well they were getting by financially (Jäckle et al. 2019b). Comparing estimates of spending with benchmark data from the UK Living Costs and Food Survey suggested that the app data collection produced broadly comparable data (Wenz et al. 2019b).

Our conclusion from the first study was that app-based data collection is promising, but that the key challenge is how to reduce barriers to participation. In the second study we focused on three aspects of app study implementation. First, we offered a browser-based alternative to the app, to accommodate those who cannot or do not want to download an app for a survey. Second, we experimentally invited sample members to the app study within an annual panel interview or by sending a letter in the post as in the first study. The expectation was that interviewers would be able to motivate and assist sample members in downloading the app, and even for those who completed the annual interview online rather than with an interviewer, the hurdle of downloading the app might seem lower if it is presented as part of a task they are already engaged in. Third, we experimented with providing participants feedback about their personal spending, as a way of motivating sample members to participate in the study and to provide complete information about their purchases. The second study was implemented in the Innovation Panel and in parallel in an online access panel. We report on the feedback experiment in a separate paper (Wenz et al. 2019a). In this paper we examine the following research questions:

1. Does introducing the app within an interview increase participation, compared to sending an invitation by post?
2. Does the effect of introducing the app within the interview vary with the mode of interview?
3. Does the in-interview invitation bring in different types of people, reducing the selectiveness of participants?
4. Does a sequential mixed-mode design, where sample members who do not use the app are offered a browser-based follow-up, increase participation?

5. Do the participants gained with the browser-based follow-up provide good quality data, compared to those using the app?
6. Does the browser-based follow-up bring in different types of people, reducing the selectiveness of participants?

2 Data

Spending Study 2 (SS2) required participants to download an app on their smartphone and use that to report all their purchases for 31 days, by entering amounts and categories of spending. The data for SS2 were collected in May to December 2018 using two samples: the *Understanding Society* Innovation Panel and an online access panel managed by Lightspeed UK. The study was implemented by Kantar Public UK and the data and documentation will be made available from the UK Data Service.

The design of the app was based on findings from in-depth qualitative interviews that explored how we could best help participants remember their daily spending (Suffield et al. 2018). The results showed that some participants are so aware of their spending that they can report fully regardless of how we ask. For others, reconstructing their day and thinking about any purchases they had made along the way was the most successful strategy, combined with prompts for categories of spending or for cash or online purchases that were more likely to be forgotten. The in-depth interviews also flagged up problems with wording, for example the term ‘spending’ was not interpreted as intended by everyone.

The app was programmed using a questionnaire app platform that Kantar Public uses for other surveys and was compatible with iPhones and Android smartphones and tablets. In one section of the app, participants were asked to enter their daily purchases, by first selecting a category and then entering the total value of the purchase, or by reporting that they had not made any purchases that day. In another section, participants were asked to enter all direct debit payments and standing orders that would come out of their accounts within 31 days. Sample members who did not use the app were invited to use a browser-based version of the spending diary instead. The design and functionality of this online diary mirrored that of the app. For documentation of SS2, including screenshots of the app and online diary, see Jäckle et al. (2019c; 2019d). The same app and online diary were implemented in the two samples, however, there were some differences between the samples in fieldwork protocols.

2.1 The Innovation Panel (IP)

The Innovation Panel is part of *Understanding Society*: the UK Household Longitudinal Study and used as a platform for methodological testing and experimentation (University of Essex. Institute for Social and Economic Research 2019a). It is a clustered and stratified probability sample of households in Great Britain, that interviews all adult household members aged 16+ annually (see Lynn 2009 for details of the sample design).

The 2018 (wave 11) interview was used as a baseline survey and all respondents were invited to SS2. Respondents were promised £1 if they completed the direct debit/standing order section, £0.50 per day on which they used the diary (whether to report purchases or a day without spending), a £10 bonus if they used the diary every day, and £3 if they completed a short debrief questionnaire at the end of the study. The incentives added up to a maximum of £29.50 and were sent to participants by post at the end of the study, in the form of Love2Shop gift cards that can be used in many high street shops.

The implementation in the IP included an experiment whereby a random half of the sample were invited to download the app within the annual interview. The other half were sent an invitation letter a couple of weeks after their interview. The treatments were allocated at the household level, so that all members of a household were treated in the same way, and stratified by allocation to mode of interview. A randomisation check, using χ^2 tests adjusted for the clustered and stratified sample design of the Innovation Panel, shows that there is a slight imbalance between respondents allocated to the two invitation treatment groups at the 5% level in terms of education, but no differences in terms of age, gender, whether in work, region of residence, frequency of using the internet, and smartphone ownership.

Regardless of the mode of invitation to the Spending Study, respondents who did not use the app were sent up to two emails reminding them to download and use the app, before being sent a letter inviting them to use the browser-based diary as an alternative way of participating in the study. The wording of the in-interview invitation to SS2 can be found in the IP11 questionnaire.¹ The schedule and content of postal/email invitations and reminders can be found in the SS2 user guide (Jäckle et al. 2019d).

¹ The IP questionnaires are available at <https://www.understandingsociety.ac.uk/documentation/innovation-panel/questionnaires>.

The IP11 interview also included an experiment with the mode of interview: a random sub-set of households were issued to face-to-face interviewers (with non-respondents followed up online), the rest were issued to web first (with non-respondents followed up by face-to-face interviewers). For more details of the IP11 design and fieldwork see Jäckle et al. (2019a).

All analyses exclude 258 respondents from a reserve refreshment sample issued in wave 11, who were invited to SS2 but for whom the logins were erroneously not activated. The analyses in this paper also exclude a single respondent who completed the IP11 interview by telephone and was not asked the self-completion module that included background questions for this study. The resulting analysis sample includes 2,637 IP11 respondents invited to SS2. Six participants used both the app and the online diary to report their purchases. They were classified as app or online diary users according to which they used more frequently.

The analyses by mode of interview (RQ2) in addition exclude 90 respondents in households with a low predicted probability of completing the IP interview online (see Jäckle et al. 2019a). These cases were all allocated to CAPI-first and are excluded, since the analysis of mode effects relies on the randomized allocation to modes. Table 1 documents the mode allocations and outcomes, excluding the 90 respondents that were not randomly allocated. Overall, 35.5% of the sample was allocated to CAPI-first and 64.5% to web-first. Nearly all respondents allocated to CAPI-first completed the survey with an interviewer (93.4%), the rest completed it online. Three-quarters of those allocated to web-first completed the survey online, the remaining quarter completed it with a CAPI interviewer. A randomisation check, using χ^2 tests adjusted for the clustered and stratified sample design of the Innovation Panel, confirms that there are no differences between respondents allocated to the two mode treatment groups at the 5% level in terms of age, gender, education, whether in work, region of residence, frequency of using the internet, and smartphone ownership.

Table 1: Mode of interview allocations and outcomes (IP11)

Mode of interview	Mode allocation			
	CAPI-first		Web-first	
	N	%	N	%
CAPI	843	93.4	418	25.4
Web	60	6.6	1,226	74.6
Total	903	100.0	1,644	100.0

Note: excludes 90 IP11 respondents invited to SS2, who were not randomly allocated to mode of interview.

2.2 *Lightspeed online access panel*

Lightspeed panellists were sent a baseline survey collecting background information on socio-demographic characteristics, mobile device access and usage, and financial behaviours, at the end of which they were invited to SS2 and asked to download the app. Those who did not download the app were immediately invited to use the browser-based version instead. At the end of the study participants were sent the same debrief questionnaire as the IP sample. The incentive scheme was in line with standard Lightspeed rewards policy: panellist could earn a maximum of 500 points (equivalent to about £5) and could exchange their incentives for vouchers or charity donations.

The implementation on the Lightspeed sample include a randomized experiment whereby part of the sample were offered feedback about their personal spending (see Jäckle et al. 2019d; Wenz et al. 2019a). However, since the offer of feedback did not affect participation rates or sample composition, for the purposes of this paper, we combine the feedback treatment groups in the analyses presented here. Two participants used both the app and the online diary. They were classified as app users, since they used the app more frequently than the online diary.

2.3 *Respondent characteristics*

To examine participation bias (RQ3, RQ6) we use data from the IP11 interview and the access panel baseline survey, collected of everyone invited to SS2. The questions in the access panel baseline survey largely mirrored those from IP11 and the question wordings are documented in the SS2 user guide (Jäckle et al. 2019d). The indicators we use include:

- Socio-demographic characteristics: gender, age, educational qualifications, and whether they did any paid work in the past week.
- Intensity of mobile device usage: measured by questions about which activities the respondent does on their smartphone.² The indicator is coded as 0, 1-9 and 10-12 and the zero category includes respondents who did not have a smartphone.
- Financial behaviours: including how frequently the respondent checks their bank balance (coded as daily, once a week, or less frequently), and whether they keep a budget.
- Spending in the last seven days: derived from questions asking how much the respondent had spent in the last seven days on a comprehensive list of 10 categories of spending and coded into spending deciles and quartiles.

² The activities included browsing websites, email, taking photos, looking at content on social media websites/apps, posting content to social media websites/apps, making purchases, online banking, installing new apps, using GPS/location-aware apps, connecting to other electronic devices via Bluetooth, playing games, and streaming videos or music.

The spending questions were not answered by 9.3% of respondents in the Innovation Panel and by 3.9% in the access panel. For these variables the analyses in Tables 5 and 7 are based on complete cases. For all other indicators the rate of missingness was less than 1%, with the exception of the Innovation Panel variables on education, frequency of checking bank balance and budgeting which were missing for around 3% of respondents. For all variables other than spending, missing observations are set to the modal categories in the corresponding sample.

We examine research questions RQ1-RQ3 using the invitation experiment implemented in the IP sample, focusing on whether or not respondents used the app to report purchases. We examine research questions RQ4 and RQ5 using data from the app and online diary from both the IP and the access panel. Since there are few IP participants who used the online diary, the final research question, RQ6, uses the access panel data only to examine differences between app and online diary users.

As the IP and the access panel differ in the composition of their samples and in the experiences the sample members have with their panel, we do not attempt to draw comparisons between the two in terms of completion rates. Instead we focus on the effects of our protocols within the samples. All analyses of the IP data take account of the clustered and stratified sample design.

3 Results

RQ1: Does introducing the app within an interview increase participation, compared to sending an invitation by post?

Introducing the app as part of the annual Innovation Panel interview significantly increased the participation rate (Table 2): 22.0% of respondents went on to use the app at least once to report a purchase, while only 11.9% of respondents invited by post went on to use the app. The results suggest that the additional participants brought in with the in-interview invitation were less committed and more likely to drop out: in the in-interview invitation group 64.6% of those who used the app at least once went on to use it in each of the four weeks, compared to 73.8% in the postal invitation group. Nonetheless the net effect of the in-interview invitation was positive: the percentage of respondents who used the app at least once in each of the four weeks remained higher with the in-interview invitation (14.2%) than the postal invitation (8.8%).

Table 2: Participation rates by invitation treatment (Innovation Panel)

	Invitation: In-Interview		Invitation: Post		P-value
	N	%	N	%	
Completed IP11 Interview	1,293	100.0	1,344	100.0	
Entered at least one purchase in app	285	22.0	160	11.9	0.000
Entered at least one purchase in each of 4 weeks	184	14.2	118	8.8	0.000

Notes: P-values from χ^2 tests.

RQ2: Does the effect of introducing the app within the interview vary with the mode of interview?

Web respondents were more likely to own smartphones than CAPI respondents (82.2% compared to 78.4%, χ^2 P-value = 0.033). Within interview modes, there were however no differences between the invitation treatment groups in smartphone ownership, as would be expected given the randomized assignment to treatments (Table 3).

Among web respondents the invitation treatment had no effect: 24.9% of smartphone owners in the in-interview invitation group and 20.8% in the postal invitation group used the app at least once to report a purchase (P=0.168). Considering all web respondents, regardless of whether or not they had a smartphone, shows no differences between invitation treatments in whether respondents used the app at least once.

Among CAPI respondents the invitation treatment however had a large effect. In the in-interview invitation group 29.5% of smartphone users used the app at least once, compared to just 7.8% in the postal invitation group (an increase of 21.7 percentage points, P<0.001). Considering the full sample of CAPI respondents, the in-interview invitation increased app usage by 17.4 percentage points.

These findings suggest that the in-interview invitation only worked if the interview was carried out by an interviewer, but not if the respondent completed the survey online. However these analyses do not account for self-selection of respondents into the mode of interview: not all respondents completed the survey in the mode to which they were randomly allocated. It is possible that respondents who completed the IP survey online, were types of people who were more likely to participate in the app study without additional encouragement.

The lower half of Table 3 replicates the previous analyses by interview mode allocation (Intention-to-Treat analysis). For both invitation treatment groups, the participation rates are lower in the Web-first group than among web respondents, and higher in the CAPI-first group than among CAPI respondents. This suggests that there are selection effects: respondents who completed the survey online were people who were more likely to participate in SS2. The Intention-to-Treat analysis

however does not allow us to draw conclusions about whether the in-interview invitation treatment is more or less effective when implemented in a CAPI or a web interview.

Table 3: Smartphone coverage and participation rates (Innovation Panel)

Mode of interview:	CAPI				Web				
	Invitation treatment:	In-Int.	Post	Δ	P-value	In-Int.	Post	Δ	P-value
Has smartphone		79.1	77.8	1.3	0.586	82.5	81.9	0.6	0.766
Used app (cond.) ¹		29.5	7.8	21.7	0.000	24.9	20.8	4.1	0.168
Used app		23.8	6.4	17.4	0.000	21.4	18.0	3.4	0.165
N		635	626			618	668		
Mode allocation:	CAPI-first				Web-first				
	Invitation treatment:	In-Int.	Post	Δ	P-value	In-Int.	Post	Δ	P-value
Has smartphone		84.1	80.0	4.2	0.178	78.9	79.9	-1.0	0.612
Used app (cond.) ¹		33.0	10.9	22.1	0.000	23.7	16.7	6.9	0.006
Used app		28.4	9.1	19.3	0.000	19.3	14.1	5.2	0.011
N		454	449			799	845		

Notes: In-Int. = In-Interview invitation. Δ = Percentage point difference between prior two columns. Cond. = conditional on smartphone ownership. ¹This row excludes 5 CAPI respondents and 11 web respondents who said they did not own a smartphone, but did use the app. P-values from χ^2 tests.

Table 4 provides estimates of the Local Average Treatment Effect, that is, the effect for those who would complete the survey in a different mode, if their mode allocation was switched. The predicted probabilities are based on a 2-stage least squares regression of the probability of using the app at least once regressed on the invitation treatment, the mode of interview, and the interaction between the two. The mode of interview and the interaction are treated as endogenous variables and instrumented with the randomized mode allocation and the interaction of the mode allocation and the invitation allocation.

Table 4: Predicted probabilities of using app (Innovation Panel)

Mode of interview:	CAPI			Web			
	Invitation treatment:	In-Int.	Post	Δ	In-Int.	Post	Δ
Used app		29.4	8.7	20.7	15.6 ^A	15.8 ^A	-0.2

Notes: N= 2,547. Predicted probabilities estimated from instrumental variable regression. Predictions sharing a letter are not significantly different. All other pairwise comparisons are significant at the 5% level. Δ = Percentage point difference between prior two columns.

The results suggest that if respondents complete the IP interview as a self-completion survey, they are equally likely to participate in a follow-on self-completion task regardless of whether they were invited to that within the web survey or by post. In contrast, if respondents complete the IP interview with an interviewer, they are unlikely to respond to the postal invitation to use the app

(8.7%) and more than three times as likely to use the app if they are invited to download it while the interviewer is present (29.4%).

RQ3: Does the in-interview invitation bring in different types of people, reducing the selectiveness of participants?

Table 5 examines participation bias: the extent and nature of differences between those who did and those who did not use the SS2 app to report purchases – and how this varies between the invitation treatment groups. The first two columns show the characteristics of all IP11 respondents in the in-interview invitation group. The third column shows how app users differ from that full sample. The numbers give the percentage point difference in the proportion of app users with that characteristic compared to the full sample. For example, 25.1% of respondents in the full sample were aged 66+, but among app users the proportion in that age group was only 6.0%, so 19.1 percentage points lower. Column 4 reports p-values from χ^2 tests for differences between app users and those who did not use the app. Columns 4 to 8 repeat this analysis for the postal invitation group. The final row summarizes the absolute average bias for the two invitation treatment groups, calculated as the sum of the absolute percentage point differences in the column above, divided by the number of rows.

The results suggest that the invitation experiment did not affect the nature or extent of participation bias. In both treatment groups there are differences in the socio-demographic characteristics of participants and non-participants: while gender is balanced, those who are younger, have higher educational qualifications, or are in work are significantly over-represented among participants. The invitation treatment altered neither the over-representation of intense smartphone users among participants, nor the over-representation of those who frequently check their bank balances. That is, there are clear differences in socio-demographic characteristics, mobile device usage, and financial behaviours that are related to participation. There does, however, not appear to be any bias in terms of what the app was designed to measure: examining the spending respondents reported in the IP11 interview shows no differences between participants and non-participants, whether the spending is grouped into quartiles (see Table 5) or percentiles (not shown).

Having examined the effects of how sample members are invited to the app study, we next examine the effects of the browser-based alternative that was offered to those who did not use the app. For this purpose we use both the Innovation Panel and the access panel data.

Table 5: Participation bias by invitation treatment group (Innovation Panel)

	Invitation: In-Interview				Invitation: Post			
	Full sample		App users – Full sample		Full sample		App users – Full sample	
	N	Col%	% points	p-value	N	Col %	% points	p-value
Female	717	55.5	2.8	0.225	737	54.8	2.0	0.548
Age 16-35	293	22.7	13.8	.	306	22.8	12.9	.
Age 36-50	296	22.9	9.7	.	311	23.1	-1.9	.
Age 51-65	380	29.4	-4.5	.	390	29.0	3.5	.
Age 66+	324	25.1	-19.1	0.000	337	25.1	-14.4	0.000
Degree	471	36.4	8.1	.	564	42.0	4.9	.
A/AS levels	183	14.2	3.7	.	170	12.6	3.0	.
GCSE/CSE level	423	32.7	-1.8	.	369	27.5	0.0	.
No educational qualification	216	16.7	-10.0	0.000	241	17.9	-7.9	0.030
In work	714	55.2	19.2	0.000	741	55.1	14.2	0.001
Activities on smartphone: 0	350	27.1	-21.8	.	397	29.5	-22.0	.
Activities on smartphone: 1-9	488	37.7	-2.3	.	491	36.5	2.2	.
Activities on smartphone: 10-12	455	35.2	24.1	0.000	456	33.9	19.8	0.000
Checks bank balance: most days	287	22.2	9.4	.	259	19.3	13.2	.
At least once a week	487	37.7	4.4	.	505	37.6	3.1	.
Less often	519	40.1	-13.8	0.000	580	43.2	-16.3	0.000
Does not keep a budget	802	62.0	-4.1	0.180	810	60.3	-1.5	0.709
Spending quartile 1	302	25.7	-2.5	.	308	25.3	-2.3	.
Spending quartile 2	298	25.4	-0.4	.	296	24.3	1.4	.
Spending quartile 3	295	25.1	-0.1	.	304	25.0	0.7	.
Spending quartile 4	280	23.8	3.0	0.628	309	25.4	0.3	0.917
Average absolute bias			8.5				7.0	

Notes: P-values from χ^2 tests.

RQ4: Does a sequential mixed-mode design, where sample members who do not use the app are offered a browser-based follow-up, increase participation?

Table 6 documents participation rates at different stages, in both the access panel and the Innovation Panel. Respondents who were invited to SS2 within the baseline interview, that is all access panel members and the IP respondents in the in-interview invitation group, were asked to report in the survey whether they had tried to download the app. In both samples about a quarter of those asked confirmed that they had successfully downloaded and logged into the app. It was not possible to identify who had actually downloaded the app, therefore we next examine who actually used the app. In the Innovation Panel most people who had said they had logged into the app went on to use it at least once. This is true both for web and CAPI respondents: within the in-interview invitation group, there is no difference between modes in the percentage who used the app at least once (χ^2 test $P=0.534$). In contrast, in the access panel there was a large drop from 24.2% who downloaded and logged into the app to 14.8% of the sample who used the app at least once. In all samples and groups, nearly everyone who opened the app entered at least one purchase.

In the access panel, 42.2% of respondents opened the online diary at least once, however only 22.8% entered at least one purchase. Nonetheless this increased the participation rate from 14.2% with only the app, to 37.0% overall. In the Innovation Panel, most respondents who opened the online diary did enter at least one purchase, however the rates were extremely low, only 3.0% overall.

These results suggest that the sequential design can more than double participation rates, but seemingly only if the alternative is also offered within the interview, as soon as respondents decline the app. We interpret these results with caution, since they come from two very different samples and so the differences in outcomes could be due to factors other than the protocols we controlled.

Table 6: Participation rates at different stages (access panel and Innovation Panel)

	Access Panel		IP: In-interview		IP: Postal		IP: All	
	N	%	N	%	N	%	N	%
Completed baseline survey	2,878	100.0	1,293	100.0	1,344	100.0	2,637	100.0
Downloaded & logged in	696	24.2	330	25.5	-	-	-	-
Accessed app at least once	427	14.8	303	23.4	162	12.1	465	17.6
... entered 1+ purchase	408	14.2	285	22.0	160	11.9	445	16.9
Accessed OD at least once	1,213	42.2	34	2.6	53	3.9	87	3.3
... entered 1+ purchase	657	22.8	29	2.2	50	3.7	79	3.0
1+ purchase in app or OD	1,065	37.0	314	24.3	210	15.6	524	19.9

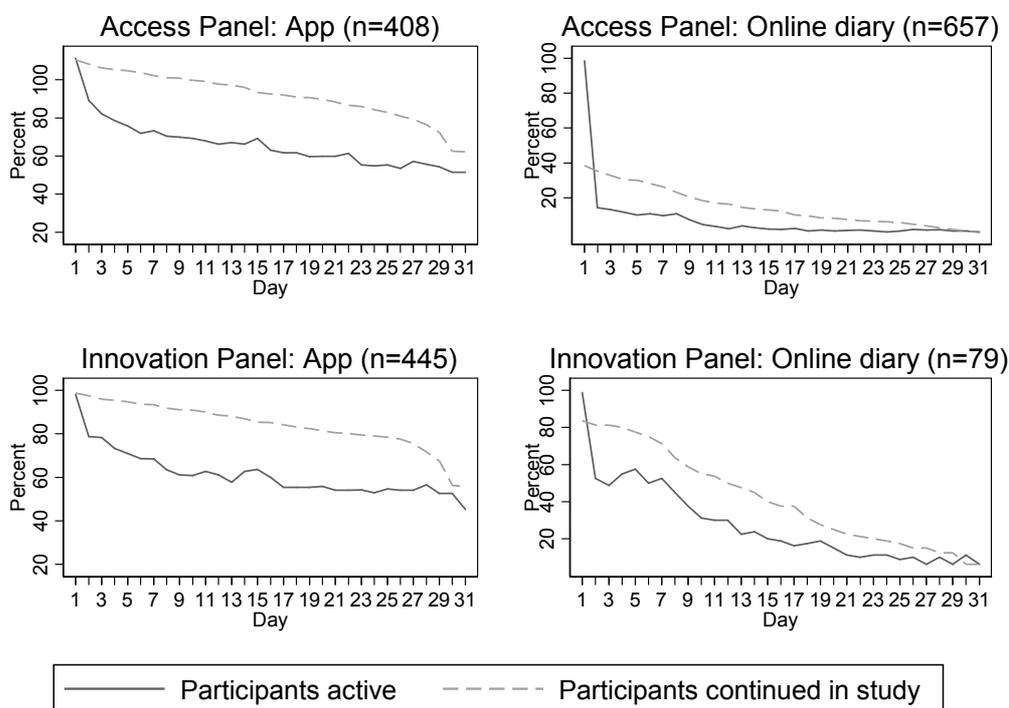
Notes: OD = online diary.

RQ5: Do the participants gained with the browser-based follow-up provide good quality data, compared to those using the app?

Figure 1 illustrates the patterns of participation over the 31 days, using the access panel and Innovation Panel data. The graphs are based on participants who reported at least one purchase. For each participant, day 1 is the day on which they first used the app or online diary. The solid lines show the proportion of participants who used the diary on a given study day (to enter a purchase, direct debit, standing order, or to report a no spend day). The dashed lines show the proportion who remained in the study and used the diary at least once on a later day.

Those who used the app maintained their participation during the study period: about 80% remained in the study past day 25 in both samples, although they did not use the app every day. In contrast, those who used the online diary did not maintain their participation over the month. In the access panel only 40% of online diary users used it on a second day and by day 25 only about 10% remained in the study. In the Innovation Panel the initial dropout was less steep, with about 80% still in the study by day 5. However, by day 25 fewer than 20% remained in the study.

Figure 1: Participation over the month



Examining the first day on which each participant used the diary suggests that, on day 1 at least, the reporting behaviours of those who used the online diary were no worse than those of app users. In the Innovation Panel, 46.8% of online diary users and 40.6% of app users entered information about direct debits and standing orders, around 77% reported at least one purchase, and those reporting

purchases made on average 2 entries. In the access panel a higher proportion of online diary users reported direct debits or standing orders (58.8% compared to 46.3% of app users), a higher proportion reported at least one purchase (87.7% compared to 77.4% of app users), and the mean number of purchases reported was similar (1.5 compared to 1.9 among app users).

These results suggest that the online diary and app might produce similar reporting behaviours and that the main challenge is how to maintain daily participation in the online diary. There are some aspects of the fieldwork protocols that differed between the app and the online diary which might in part be responsible for the differences in dropout patterns. App users received a daily push notification sent out at 8pm from within the app, reminding them to report their purchases for that day. For the online diary we would have liked to send out a similar reminder by email. Emailing everyone every day seemed excessive and ideally we would have wanted to email only those who had not used the app by 8pm that day, however such customized mailings could not be automated. Although all participants received a weekly email summarizing the incentives they had earned so far and encouraging them to continue participating, it seems the weekly email was not enough to sustain daily participation.

RQ6: Does the browser-based follow-up bring in different types of people, reducing the selectiveness of participants?

Table 7 examines differences between participants and non-participants, and how the extent and nature of bias changes when those who used the online diary are added to the participant pool. The analyses are based on the access panel data only, since the number of online diary users in the Innovation Panel is too small to examine changes in sample composition. Columns 1 and 2 show the characteristics of respondents who completed the baseline questionnaire. For each characteristic, column 3 documents the percentage point difference in the proportion of app users with that characteristic. Column 4 reports p-values from χ^2 tests of differences between app users and those who did not use the app. Columns 5 and 6 repeat this analysis, comparing those who used either the app or the online diary with those who did not participate in either way.

The selectiveness of app users compared to all baseline respondents mirrors the pattern in the Innovation Panel (Table 5 above). Once online diary users are added to the participant pool, the extent of bias decreases: the average absolute bias for the variables examined more than halves, from 6.4 to 2.8. However the biases remain significant. That is, offering the browser-based alternative reduces but does not eliminate participation bias.

Table 7: Participation bias of app users and combined app and online diary users (access panel)

	Full sample		App users – Full sample		App/OD users – Full sample	
	N	Col %	% points	P-value	%pts	P-value
Female	1,998	69.4	3.1	0.139	1.5	0.190
Age 16-35	986	34.3	1.3	.	1.1	.
Age 36-50	777	27.0	10.3	.	4.4	.
Age 51+	1,115	38.7	-11.5	0.000	-5.5	0.000
Degree	1,205	41.9	3.7	.	2.0	.
A/AS levels	740	25.7	2.7	.	1.0	.
GCSE/CSE level	795	27.6	-4.8	.	-1.6	.
No educational qualification	138	4.8	-1.6	0.023	-1.3	0.019
In work	1,901	66.1	7.7	0.000	3.2	0.005
Activities on smartphone: 0	458	15.9	-14.2	.	-4.6	.
Activities on smartphone: 1-9	1,268	44.1	-5.3	.	-1.6	.
Activities on smartphone: 10-12	1,152	40.0	19.5	0.000	6.3	0.000
Checks bank balance: most days	945	32.8	12.0	.	4.0	.
At least once a week	1,165	40.5	-0.3	.	1.8	.
Less often	768	26.7	-11.7	0.000	-5.7	0.000
Does not keep a budget	588	20.4	3.8	0.038	0.2	0.817
Spending quartile 1	692	25.0	-3.3	.	-3.2	.
Spending quartile 2	712	25.7	-3.8	.	-1.5	.
Spending quartile 3	675	24.4	3.6	.	1.3	.
Spending quartile 4	688	24.9	3.6	0.023	3.4	0.001
Average absolute bias			6.4		2.8	

Notes: OD = online diary. P-values from χ^2 tests.

4 Conclusion

This paper provides novel evidence on how protocols used to implement app-based data collection can affect participation rates and biases.

Using experimental data we show that inviting respondents to the app-based task within a CAPI interview more than triples participation rates compared to sending an invitation letter by post. In contrast, inviting respondents to download the app within a web survey offers no advantage over sending the invitation by post.

Offering a browser-based alternative to those who do not use the app can more than double participation rates. However this appears to only be effective if the invitation to the app study is embedded in a survey and if the alternative is offered as soon as the app is declined. The browser-based alternative brings different types of people into the participant pool, reducing non-participation bias. A key challenge with the browser-based alternative, however, is how to maintain participant engagement throughout the study period: while the app can be set to push out daily reminders, it is less clear how to implement comparable reminders for a browser-based instrument.

References

- Jäckle, A., T. Al Baghal, J. Burton, O. Kaminska, and P. Lynn (2019a) Understanding Society: The UK Household Longitudinal Study Innovation Panel, Waves 1-11, User Manual. Colchester: University of Essex.
- Jäckle, A., J. Burton, M.P. Couper, and C. Lessof (2019b) 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. *Survey Research Methods*, 13(1):23-44.
- Jäckle, A., Jonathan Burton, A. Wenz, and B. Read (2019c) Understanding Society: the UK Household Longitudinal Study Spending Study 2. User Guide – Appendix A: App and Online Diary Screenshots. Colchester: University of Essex.
- Jäckle, A., Jonathan Burton, A. Wenz, B. Read, T. Hanson, and D. Xu (2019d) Understanding Society: The UK Household Longitudinal Study Spending Study 2 User Guide. Colchester: University of Essex.
- Lessof, C., A. Jäckle, M.P. Couper, and T.F. Crossley (2019) Adherence to Protocol in a Mobile App Study Collecting Photographs of Receipts: Participant Behaviours and Predictors of Adherence. Unpublished manuscript.
- Lynn, P. (2009) Sample Design for Understanding Society. Understanding Society Working Paper 2009-01. Colchester: University of Essex.
- Read, B. (2019) Respondent Burden in a Mobile App: Evidence from a Shopping Receipt Scanning Study. *Survey Research Methods*, 13(1):45-71.
- Suffield, M., H. Hasbrouck, A. Coulter, A. Jäckle, J. Burton, T.F. Crossley, M.P. Couper, and C. Lessof (2018) Understanding How People Think About Their Daily Spending. Understanding Society Working Paper 2018-02. Colchester: University of Essex.
- University of Essex. Institute for Social and Economic Research (2019a) Understanding Society: Innovation Panel, Waves 1-11, 2008-2018. [data collection]. 10th Edition. UK Data Service. SN: 6849 <http://doi.org/10.5255/UKDA-SN-6849-11>.
- (2019b) Understanding Society: Spending Study 1, 2016. [data collection]. UK Data Service. SN: 8348. Dataset. <http://doi:10.5255/UKDA-SN-8348>.
- Wenz, A., A. Jäckle, J. Burton, and M.P. Couper (2019a) The Effects of Feedback on Participation and Reporting in Mobile Data Collection. Unpublished manuscript.
- Wenz, A., A. Jäckle, J. Burton, M.P. Couper, and B. Read (2019b) Quality of Expenditure Data Collected with a Receipt Scanning App in a Probability Household Panel. Unpublished manuscript.