Understanding Society Working Paper Series

No. 2026 – 01

January 2026

How Panel Members React to Multiple Requests for Additional Data over Time

Jasmine Mitchell and Annette Jäckle

Institute for Social and Economic Research, University of Essex

Non-Technical Summary

Enhancing survey panel data by linking to existing digital data or using methods that do not rely on self-reporting is increasingly of interest. All such methods, however, still rely on the respondent's willingness to participate in the respective data collection task. As individuals are increasingly being asked to do more than just answer questions in surveys, the overarching question is how to design and implement a set of tasks using different ways to gather data on different concepts, using different methods, in a way that sample members will cooperate.

This paper is the first examination of how panel members react to multiple requests for different types of additional data over time. To investigate this, we used data from 14 additional tasks, implemented in the *Understanding Society* Innovation Panel over a 10-year period, for which respondents had to use several mobile apps, supply bio-measures, consent to data linkages, and participate in monthly mini-surveys and time diaries.

The results indicate there are high rates of churn in the sample, with individuals flowing in (due to household joiners and refreshment samples), tending to participate in just under half the tasks they were invited to, and then possibly leaving the sample (mainly due to attrition). There are no clear patterns of participation across task type, topic, and whether they are incentivised; the more tasks respondents were invited to, the more different types they participated in. However, the more tasks individuals were invited to, the less likely they were to participate in the later tasks and in the later annual interviews. Therefore, repeatedly inviting individuals to participate in additional tasks has a small detrimental effect on survey panels.

How Panel Members React to Multiple Requests for Additional Data over Time

Jasmine Mitchell and Annette Jäckle

Institute for Social and Economic Research, University of Essex

Abstract: Over the years, the *Understanding Society* Innovation Panel has asked respondents to complete various additional data collection tasks. For example, providing hair or blood samples, measuring fingers, consenting to link administrative data, using mobile apps, and completing time use diaries. We examine the cumulative effects of these additional tasks on participation in later tasks and annual interviews. We find no systematic patterns of participation in the multi-modal set of tasks. However, the more tasks individuals were invited to, the less likely they were to participate in the later tasks and the later annual interviews. Therefore, repeatedly inviting individuals to participate in additional tasks has a small detrimental effect on survey panels.

Keywords: additional data tasks, panel survey, non-response, prior survey requests, consent to data linkage, mobile app data collection, biomarker collection.

JEL classification: C80, C83.

Acknowledgements: This research was funded by Economic and Social Research Council (ESRC) grants for Understanding Society (ES/K005146/1, ES/N00812X/1, ES/S007253/1, ES/T002611/1) *Understanding Society* is an initiative funded by the ESRC and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Verian. The research data are distributed by the UK Data Service (study numbers: 8990, 9065, 6849, 8749, and 8909).

Corresponding Author: Jasmine Mitchell, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom, jasmine.mitchell@essex.ac.uk.

1. Introduction

Survey data collection methods have evolved with the introduction and rise of new technologies (Link et al., 2014). Methods involving technologies such as global positioning systems, bio-measures, sensors, and mobile apps, amongst other new forms of data collection, have become increasingly prevalent among researchers to collect data for surveys (Couper, 2017; Struminskaya et al., 2020).

These newer data collection methods can provide more accurate and detailed data than self-reports, as well as information not known by respondents. They do, however, come with a host of challenges with both implementing the data collection and the potentially selected nature of participants (for example, Annette Jäckle et al., 2023). As with surveys, it is important to investigate and research the Total Survey Error to understand potential biases in data collected with these additional tasks and how best to implement such technologies (Groves et al., 2009).

The adoption of these new technologies means individuals are increasingly being asked to do more than just answer questions in surveys. Therefore, this begs the question of how we design and implement a multi-modal data collection system to gather data on different concepts, in a way that sample members will cooperate.

As reviewed throughout this paper, there is an increasing amount of research on new methods of data collection and their implementation, albeit some more researched than others. However, the growing body of research has mostly investigated new data collection methods independently from one another. Thus far, research on individuals' participation in numerous multi-modal tasks has been neglected. This paper is the first examination of how panel members react to multiple requests for different types of additional data over time.

The focus of this paper is on 14 additional tasks that were implemented throughout 10 waves of a nationally representative longitudinal panel in Great Britain (*Understanding Society* Innovation Panel; IP), over a period of 10 years. The IP is a test-bed panel for innovative methods of data collection which collect multi-modal data, including consent to data linkage, various app studies, monthly mini-surveys, time diaries, bio-measure and finger measurement studies. We use these data to examine the following research questions:

1. What are the patterns of participation across additional data collection tasks? For example, are there respondents who participate in one type of task (e.g. linkage consents) but not others (e.g. mobile app studies)?
2. Do respondents become less likely to participate in additional tasks the more tasks they are asked to do?
3. Do respondents become less likely to participate in the annual interview the more additional tasks they are asked to participate in?

2. Background

2.1 Participation rates in different types of additional tasks

Appendix A summarises examples of prior studies using data collection methods of the types included in our analyses, and the participation rates they achieved.

Several methods tend to elicit participation rates over 50%. These include bio-measures (Fitzsimons et al., 2020; Kearney et al., 2011; Sakshaug et al., 2014), data linkage consent (Burton, Couper, et al., 2024; Mostafa & Wiggins, 2018), consent to ask questions via short messaging service (SMS) (Cooke et al., 2003; Annette Jäckle et al., 2023), time-use diaries (Abraham et al., 2006; Ingen et al., 2009; Morris et al., 2016; Office for National Statistics, 2003), finger measurement studies (Allum et al., 2014), and mini-monthly surveys (A Jäckle, J Burton, M.P Couper, Vine. J, et al., 2023). Unlike the other methods, mobile apps tend to result in participation rates lower than 50%. This is potentially due to the number of steps individuals are presented with before they can use an app, such as having a compatible device (Jäckle et al., 2019; Torous et al., 2020).

That said, participation rates vary within tasks as well. Findings for tasks with a more extensive body of research are inconsistent, there is no consensus of the participation or consent rates achieved (Burton, Couper, et al., 2024; Sakshaug et al., 2012). Specifically for mobile apps, response rates vary dramatically. For example, Annette Jäckle et al. (2023) report on three different mobile app studies with participation rates of 13%, 17/18%, and 45%. This suggests that participation rates vary substantially with app characteristics.

Overall, the evidence suggests that participation rates for additional tasks differ considerably, both between and within task types. These variations may be related to task factors, such as required time, cognitive demands, or type of data collected, as well as respondent characteristics, such as motivation, familiarity, and prior survey experience.

Related to RQ1, these findings imply that there is likely little consistency of participation between and within additional data collection tasks. Respondents may participate in some tasks, but not others, and this is also the case for repeated tasks of the same type. Research on some of these methods remains scarce. In addition, existing studies investigate additional tasks independently rather than cumulatively. This highlights the importance of examining the patterns of participation between and within various additional data collection tasks.

2.2 Impact of previous requests on response rates in later surveys

Overall, most panel studies find a negative but small effect of the number of previous survey or task invitations on participation in later surveys.

As cited by Yan and Williams' (2022) response burden conceptual framework, an important factor that contributes to unit non-response and attrition is prior survey experience and frequency of interviews. As the number of surveys and other data collection tasks increase, survey fatigue and survey saturation may occur (Groves & Couper, 2012). Additionally, some respondents meet new data collection technologies with privacy and data security concerns that could further affect participation or attrition (Keusch et al., 2019; Revilla et al., 2019). Although some studies have found that additional burden can decrease the likelihood of continued participation in panels, others have shown that it does not affect dropout or attrition (Lynn, 2014; Sharp & Frankel, 1983).

Many studies that identified negative effects only found a small decrease in participation in later surveys. Eisnecker and Kroh (2016) used refreshment samples from the German Socio-Economic Panel and found that attrition rates between waves one and two increased by only 1.5 percentage points for those asked for consent to employment and benefits records compared to those that were not. Similarly, Eggleston (2024) analysed survey response to the 2020 Census for the United States and found that households invited to respond to the American Community Survey around a year before the census had a two percentage point lower response rate to the census than the national census sample. However, this did decrease further to 15 percentage points for those who were sampled for the American Community Survey a few months before the census. Similar results were found for individuals who were previously invited to the Current Population Survey. This suggests that the time between being invited to surveys affects participation in later surveys, and that increasing the time between survey requests may reduce this negative effect.

Studies that used pre-established panel members found similar results. For example, Trappmann et al. (2023) asked three quarters of the German PASS panel to participate in a research app and provide passive sensor data over six months. They found panel retention decreased by three percentage points compared to the control group not invited to the study. Interestingly, they observed that the effect diminished in the subsequent waves after the invitation. Similar results are found in Great Britain. A Jäckle, J Burton, M.P Couper, Vine. J, et al. (2023) investigated if being invited to monthly life event surveys affected attrition in the annual IP interviews. They found that 65% of those not invited to the additional task and 64% of those that were invited participated in the next annual interview, suggesting that inviting individuals to the monthly life events surveys had little effect on attrition in later interviews.

Based on prior evidence related to RQ2 and RQ3, we expect that respondents will be less likely to participate in additional tasks and annual interviews the more additional tasks they are asked to participate in. On the one hand this effect is expected to be small, meaning that the benefits of new and potentially more reliable data may outweigh the costs of a small increase in attrition. On the other hand, previous research has primarily examined the impact of a single survey request on panel participation, rather than the cumulative effects of multiple tasks. As the number of additional tasks increases, the negative effects on attrition could become substantially larger.

2.3 Participation in sets of multi-modal tasks

Although no research has been conducted on sets of multi-modal tasks, some studies have investigated hypothetical willingness to participate in different types of tasks. Revilla et al. (2019) analysed self-reported willingness to complete several different tasks in a web survey using the *Netquest* opt-in panel in Spain. The tasks included passive measurement on devices they already use or on new devices, self-reporting results from using measurement devices and the collection of bio-measures. They found that willingness varied across the different types of tasks and was higher for those in which participants have control over the reporting of results rather than passive data collection tasks, stating this is due to privacy and trust concerns. Whilst this research provides an insight into participation across multiple different data collection methods, it is an opt-in panel using sample members who are generally more cooperative and solely investigates stated willingness, not whether the respondents actually participate in the tasks.

Perhaps the closest research on participation in multiple survey requests comes from joint modelling of consents for data linkage. Jenkins et al. (2006) examined three data linkage consent questions (benefit and tax credit administrative records, national insurance numbers, and data from employers) and found differing consent rates and biases across the different consent types. They also found positive correlations in unobservable factors influencing consent, indicating a latent consent propensity. Similarly, Mostafa (2016) used joint modelling to analyse correlates of consent to multiple data linkages in the Millennium Cohort Study. They reported varied consent rates across domains (education, health and economic records data linked) as well as correlations within individuals. Respondents who dropped out from the survey at least once were less likely to consent to all data linkages, further suggesting that there is a latent propensity to cooperate.

However, these studies used surveys that asked for multiple consents in the same interview. Research examining consent data across multiple time points have provided evidence for weak correlations between unobservables. Mostafa and Wiggins (2018) used the Millennium Cohort Study to analyse main respondents’ consents to linking their children’s health records in three different waves. Results indicated that correlations between unobserved parts of consent outcomes over time were low, with respondents’ choices seemingly driven by current circumstances and interviewer effects. This suggests that stable latent characteristics are not the main drivers of consent, but the circumstances of the respondent at the time.

Further supporting our expectations for RQ1, these studies suggest that participation in different data collection tasks will vary. This may depend on factors like control over the task, privacy concerns, and individual circumstances. The variation in participation behaviours across tasks and over time are likely context-dependent and may differ based on current circumstances of the individual.

3. Data

We use data from the *Understanding Society* Innovation Panel annual interviews as well as from the additional tasks that respondents were invited to complete.

3.1 *Understanding Society* Innovation Panel wave 1 to wave 16

The Innovation Panel (University of Essex, 2024) is part of *Understanding Society*: The UK Household Longitudinal Study. The design and implementation of the Innovation Panel (IP) is based on the main *Understanding Society* survey. For each household in the sample, all members aged 16 and over are interviewed annually on several topics, including modules on health, family, employment, education, and socio-economic status. Household members become eligible for the adult interviews once they turn 16.

There are approximately 1,500 households from Great Britain in the sample. The IP uses clustered and stratified probability sampling and includes refreshment samples of about 500 respondent households each in waves four, seven, 10, 11, and 14. For the current research, data from wave six (2013) to wave 16 (2023) are included. Fieldwork takes place over the summer for about five months.

Most waves use an experimental mixed mode design, whereby a random two-thirds of households are issued to web-first, with a face-to-face then telephone follow up for non-respondents. The other third of households are issued to face-to-face interviews first, with non-respondents followed by web then telephone. In IP12 one third of households were issued to nurses, one third to face-to-face interviewers, and one third to web-first. Due to the Covid-19 pandemic, all households in IP13 and IP14 were issued to web-first, with non-respondents followed up by telephone interviewers. For information of which modes where used in which waves, see Institute for Social and Economic Research (2024).

Household response rates in the annual interviews in waves one to 16 range between 59.0% and 84.7% (AAPOR RR6, The American Association for Public Opinion Research, 2023). Individual response rates, based on all eligible members in sampled households, range from 76.5% to 88.9%. For more information on response rates in the annual interviews see Institute for Social and Economic Research (2024).

3.2 The additional tasks

The following provides an overview of the additional tasks that we include in our analyses. Some of the tasks included experiments which we do not examine here since they have been reported in previous papers that we reference below. For further details on the additional tasks see Appendix B and C.

3.2.1 Consent tasks

For each of the consent questions, respondents were coded as 'participant' if they consented and 'non-participant' if they did not consent or did not answer the question.

- Financial Conduct Authority (FCA) consent: In IP9, respondents were asked if they consented to link their survey answers to FCA data containing information on credit accounts and credit rating scores.
- HM Revenue and Customs (HMRC) consent: In IP11, respondents were asked permission to link their HMRC records containing information on employment, income, National Insurance contributions, and tax credits to their survey answers (Jäckle et al., 2024).

- SMS survey consent: In IP13, all IP respondents who stated they use a mobile phone were asked for consent to occasionally receive SMS text messages containing survey questions (Vine et al., 2023).

3.2.2 App tasks

- Spending Study 1 (SS1, University of Essex, 2022a): Following the IP9 interviews, sample members were asked to upload shopping receipts or report spending directly in an app for a month. Sample members were coded as participants if they used the app at least once (Jäckle et al., 2019).
- Spending Study 2 (SS2, University of Essex, 2022b): IP11 respondents were asked to keep a diary on an app for one month reporting all expenditures. Those who provided at least one spending amount were coded as participants (Jäckle et al., 2022).
- Wellbeing app: IP13 respondents were asked to download an app and, every evening for 14 days, answer a set of questions on their interactions with loved ones, stressors, and mood. Sample members were coded as participated if there was any data from the daily app questions (A Jäckle, J Burton, M.P Couper, & B Perelli-Harris, 2023).
- Body Volume Index (BVI) app (data released with the IP15 annual interview data, University of Essex, 2024): IP15 respondents were asked to take two pictures of their body using their device camera and answer profile questions. If measurements based on the photos was obtained, sample members were coded as participants (Vine et al., 2023).
- Spatial Navigation game (data released with the IP16 annual interview data, University of Essex, 2024): IP16 respondents were invited to use a cognition game app where they participated in a series of increasingly difficult maritime-themed levels. Respondents were coded as a participant if they completed at least one level (Burton, Jäckle, et al., 2024).

3.2.3 Bio-measures

- Hair sample: Respondents were provided kits with instructions for themselves, someone else, or a professional nurse to take a hair sample. Respondents were coded as a participant if one of their hair hormone level measurements was in the data, indicating they supplied a useable hair sample (Al Baghal et al., 2025).
- Dried blood sample: Respondents were asked to use a lancet to provide blood from their finger, via the participant doing this themselves, someone else, or by a professional nurse. If a dried blood biomarker measure exists in the data, individuals were coded as participated (Benzeval et al., 2023).
- Pre-interview blood measure: The advance letter requested all individuals eligible for IP12 to take their blood pressure prior to the annual interview and then provide those measurements in the interview. IP respondents were coded as participants if they said they had their blood pressure taken before the interview (Al Baghal et al., 2025).

3.2.4 Mini monthly life event surveys

Between IP11 and IP13 sample members were invited to complete a survey each month. If they had not experienced any of a list of events, the survey ended after the first question. If they had they were asked follow-up questions about the event. Sample members were coded as participants if they completed at least one survey (A Jäckle, J Burton, M.P Couper, Vine. J, et al., 2023).

3.2.5 Time-use diary

IP7 respondents were asked to keep a time diary on two separate days recording how they spent their 24-hour day in 10-minute periods. IP respondents were coded as participants if they provided activity codes for at least one diary day (Institute for Social and Economic Research, 2024).

3.2.6 Finger Measurement Study

In IP6, all respondents were asked to provide measurements of their second and fourth finger. IP respondents were coded as participants if there was data for at least one finger measurement (Institute for Social and Economic Research, 2024).

3.3 Task outcomes and codes

We compute 16 outcome codes for the 5,812 IP respondents who were eligible for at least one task, one for each of the additional tasks. The outcomes were coded as participant, non-participant, or ineligible. Ineligibility is defined as a panel member having certain characteristics, such as being under 16, part of a later refreshment sample, etc, which prevented them from being invited to the task. As the tasks had different eligibility conditions, those not eligible were further coded to differentiate reasons (see Appendix D). Appendix E documents the cases that are in the survey data but were not eligible for any of the tasks, and are therefore excluded from our analysis sample.

3.4 Outcome variables

We focus on two outcome variables. ***Participation in a given task*** was collapsed from the full participation and ineligibility coding frame described in Section 3.3 and documented in Appendix D. For each task, sample members were coded as 2 if they participated in that task, 1 if they were eligible for the task but did not participate, and 0 if they were not eligible. For the second research question, observations where the sample member was not eligible for the task were dropped and the outcome was recoded as 1 (participant) or 0 (non-participant).

Participation in the annual interview was coded as 1 if the sample member provided a full interview and 0 if they were eligible for the interview but did not participate. Non-respondents include full household refusals and non-contacts and other reasons for non-response such as language barriers, frailty, illness and absence during the fieldwork period.

3.5 Independent variables

The main independent variable is the ***number of previous tasks invited to***. The variable measures how many tasks panel members had been invited to prior to the current task or annual interview and therefore ranges from zero to 13. We include this as a continuous variable in our analyse, since the aim is to investigate the average effect of inviting a respondent to an additional task on the outcomes.

As the characteristics of the task itself will influence whether sample members participate, all inferential analyse concerning participation in the tasks include a binary dummy variable to ***control for what the current task was***.

The year and wave the annual interviews were implemented in also potentially influence participation in the annual interviews. Therefore, ***interview year*** binary dummy indicators are included in any analysis of participation in the annual interviews.

Due to refreshment samples in waves four, seven, 10, 11 and 14, some individuals were ineligible for certain tasks and annual interviews, before they flowed into the sample. Therefore, ***sample origin*** binary dummy variables are included in the analyse.

The analyse also includes ***controls for the task-related experiments*** (documented in Section 3.2 and Institute for Social and Economic Research, 2024). After estimating the full model, Wald tests were conducted for each task-related experiment to test whether all coefficients related to different treatments were jointly equal to zero. Experiments which did not provide significant F-test coefficients at the 0.05 level were removed. The parsimonious model includes controls for all significant task-related experimental treatments as binary dummy variables. For each of these indicators, panel members who were not in the sample that year were coded as zero. See Appendix F for more information on the task experiments as covariates.

The analyses of participation in the annual interviews also includes ***controls for experiments related to participation in the interviews***. The same Wald test procedure was used to select which experiments to control for in the parsimonious model. All individuals not included in the sample for that year were coded as zero. See Appendix G for more information on the interview experiments as covariates.

4. Methods

To examine RQ1, the patterns of participation across the different additional tasks, we conduct descriptive analyses. First, we use sequence analysis to group sample members with similar trajectories of participation status over the 14 tasks. Using the *sqindexplot* command in Stata, we generate a sequence index plot using the person identifier, the task identifier, and the outcome of whether individuals participated, did not participate, or were ineligible for the task. This sequence analysis provides a first visual glimpse into the frequency of different patterns of participation across sample members over the 14 tasks. It also visualises the inflows (due to refreshment samples) and outflows (due to survey non-response or task eligibility criteria) from the set of eligible sample members for each particular task.

We then compute the percentage of eligible tasks each respondent participated in and then average over all sample members to calculate mean participation rates. We plot these participation rates by number of tasks the sample member was eligible for, to get a sense of the relationship between number of tasks and participation rates in the raw data.

Next, we classify each sample member as to their personal pattern of participation across different types of tasks. We do this in three ways: 1) classifying tasks by whether they occurred within the interview (Finger Measurement Study, FCA consent, HMRC consent, and SMS text consent), outside the interview with supplies provided (Time Diary, pre-interview blood pressure, Life Events Study, and bio-measures), and mobile apps (SS1, SS2, Wellbeing app, Body Volume app, and Spatial Navigation Game); 2) classifying tasks by topic, whether they were related to health, finance, or life events; and 3) classifying tasks by whether we offered respondents additional incentives if they participated in the task. For each of these three classifications, we create an indicator of whether the sample member participated in only one of the types of tasks or in multiple ones, and which. We again plot the distributions of participation patterns by number of tasks the respondent was eligible for.

Finally, we focus on the sub-set of sample members who were eligible for all 14 tasks. We use the *xtdescribe* command in Stata to describe missingness patterns, using the person and task identifiers to uniquely identify each observation in the data. From this, we create an indicator that summarises each sample member's sequence of participation and non-participation across the 14 tasks. We then tabulate the frequency distributions of these sequences to examine whether there are any prevalent patterns.

Our analyses, of whether the number or tasks an individual has been invited to affects their participation in the following task (RQ2) or annual interview (RQ3), are potentially affected by selection bias. For most tasks, sample members must have participated in the annual interview that year to be eligible for the task. Therefore, those invited to more tasks are those who participated in more annual interviews. Conversely, individuals with a high number of previous task invitations are those who have not dropped out of the panel. That is, sample members invited to more tasks are potentially more cooperative. The characteristics of a sample member that make them more likely to stay in the panel might also make them more likely to participate in a given task. Such omitted variables would bias estimated effects of the number of prior task invitations on participation (Gerber & Green, 2012). We use several methods to account for such potential selection biases.

For RQ2 we use a respondent-task level dataset, which includes 31,728 observations for the 5,812 sample members who were eligible for at least one task. We estimate a series of models regressing participation in additional tasks on the number of prior tasks invited to: unweighted and weighted linear probability, logit, fixed-effects, controlling for time-varying cooperativeness, and instrumental variable models. Equation (1) is our baseline linear probability model:

$$Y_{it} = \beta_0 + \beta_1 N_{it} + \beta_2 T_t + \beta_3 S_i + \beta_4 X'_{it} + \epsilon_{it} \tag{1}$$

where Y_{it} represents the binary indicator of participation for individual *i* in task *t*. The key independent variable, N_{it}, is the number of previous tasks individual *i* was invited to prior to task *t*. The model includes controls for what the current task is (T_t), the sample origin of the individual (S_i), and a vector of controls for task-related experimental treatments that individual *i* has been exposed to up to and including the invitation to task *t* (X'_{it}). The error term, ϵ_{it}, captures unobserved factors that affect participation of individual *i* in task *t*.

As robustness checks we also estimate logit models and find that the results are similar. We present the results from the linear probability models, as they are more straightforward to interpret: the coefficient of interest (β_1) is an estimate of the percentage point change in the probability of participation associated with a one-unit change in the number of tasks the individual was previously invited to.

We then test additional specifications that each use a different method to account for potential selection effects. The first method is to use non-response weights, to account for differences in observable characteristics between those who remain in the panel and those who drop out. For the weighted model, we estimate equation (1) using the cross-sectional non-response weights supplied with the corresponding annual interview data. The weight uses the inverse response probabilities and is multiplied by the issue weight. For details on how the weights are created see Institute for Social and Economic Research (2023). Note that cases where the individuals were invited to a task but did not participate in the corresponding annual interview were dropped from the weighted analyses as they did not possess a response weight. A regression including these cases, using enumerated weights, yielded similar results (results not included).

The second method to account for potential selection effects is to estimate a fixed-effects model, based on within-individual variation in the number of prior tasks and task participation. This method nets out the effects of any fixed characteristics of the respondent that might influence participation. As indicated by equation (2), the model again regresses the probability of participation for each individual *i* at task *t* (Y_{it}) on the number of previous tasks the individual was invited to (N_{it}). The model includes a control for the task (T_t) and the fixed-effect for the individual (α_i) reflecting the fixed unobserved heterogeneity.

$$Y_{it} = \beta_1 N_{it} + \beta_2 T_t + \alpha_i + \varepsilon_{it} \tag{2}$$

To test the appropriateness of using a fixed-effects model, we conducted Hausman tests. These test the null hypothesis that α_i is uncorrelated with $\beta_1 N_{it}$ and $\beta_2 T_t$ (Verbeek, 2000). Two Hausman tests, one comparing a fixed-effects model and random-effects model and one comparing a fixed-effects and pooled model, were implemented. The results suggested a significant difference for both of the alternative models (both tests $\chi 2 = 471.29$, $p < 0.001$), indicating strong evidence in favour of using fixed-effects modelling.

Since the fixed-effects model assumes that cooperativeness is fixed within the individual, we estimate an alternative model allowing cooperativeness to vary over time. We use the percentage of questions the respondent did not answer in the corresponding annual interview as a time-varying indicator of cooperativeness. This model also controls for the mode of interview, since item non-response rates are higher in web than CAPI and CATI. The model is fitted using the equation

$$Y_{it} = \beta_0 + \beta_1 N_{it} + \beta_2 T_t + \beta_3 S_i + \beta_4 X'_{it} + \beta_5 I_{it} + B_6 M_{it} + \epsilon_{it} \tag{3}$$

where I_{it} denotes item non-response and M_{it} represents mode. As individuals who did not participate in the annual interview did not possess an item non-response indicator, they were removed from the regression.

The last method to account for potential selection effects is an instrumental variable regression, where the number of previous tasks invited to is treated as an endogenous variable and sample origin is used as an instrument. The association between the two variables is strong. Respondents from the original wave one and wave four refreshment samples were on average invited to 14 tasks, those from the wave seven refreshment sample to 13, those from the wave 10 and 11 refreshment samples to 10, and those from the wave 14 refreshment sample to two. A two-stage least-squares regression is used (Verbeek, 2000), with the first stage isolating the exogenous variation in the endogenous variable, as shown in equation (4). This regresses the endogenous number of previous tasks invited to (N_{it}) on the sample origin (S_i). As with the other models, controls for task and task experiments are also included. The predicted values $\widehat{N}$ from the first stage are then used instead of the endogenous variable (N_{it}) in the second stage equation (5).

$$N_{it} = \alpha_0 + \alpha_z S_i + \alpha_2 T_t + \alpha_3 X'_{it} + \epsilon_{1it} \tag{4}$$

$$Y_{it} = \beta_0 + \beta_1 \widehat{N} + \beta_2 T_t + \beta_3 X'_{it} + \epsilon_{2it} \tag{5}$$

To test the suitability of sample origin as an instrument, we used the Montiel-Pflueger robust weak instrument test. The test uses an effective F-statistic that is a transformation of the first stage F-statistic from the regression on the instrument, adjusting the F-statistic so that it remains valid even when errors are correlated. The tested instrument is considered strong if the two-stage least squares is substantially less biased than the ordinary least squares regression. Olea and Pflueger (2013) derive a worst-case bias benchmark which is where the instrument is effectively uninformative, with the first and second stage error being perfectly correlated. The critical values for each worst-case bias level inform how large the effective F-statistic must be to treat the instruments as sufficiently strong. For the current research, an effective F-statistic of 948.46, surpassed critical values at a 5% confidence level across various levels of worst-case bias, indicating the null hypothesis that the instrument is weak can be rejected, and that sample origin is a sufficient instrument.

For RQ3 we use a respondent-task level dataset, including all IP members who were eligible for at least one annual interview between wave one and wave 16. This totalled to 61,178 observations for 9,869 sample members. We use similar models as for RQ2, but regressing participation in annual interviews (rather than participation in tasks) on the number of prior tasks invited to: unweighted and weighted linear probability, fixed-effects, and instrumental variable models. Since we cannot calculate item non-response rates for non-respondents, we did not use the model allowing for time-varying cooperativeness. We include controls for the different experiments that were conducted in the annual interviews as covariates, rather than controls for the experiments in the additional tasks.

For the weighted model, cases which did not participate in the annual interview did not possess response weights. Therefore, enumeration weights were used instead. For the fixed-effects model, the Hausman tests indicated the use of a fixed-effects regression over a random model or over a pooled model (both tests $\chi2 = 20{,}394.74$, $p < 0.001$). For the instrumental variable regression, the effective F-statistic of 725.838 was large compared to the benchmarks, indicating that the sample origin was again a strong instrument.

All analyses were conducted using *Stata Version 18* and account for the clustered and stratified sample design. All graphs, excluding the sequence analysis, were created using version 4.4.2 in *RStudio*.

5. Results

5.1 What are the patterns of participation across additional data collection tasks?

The sequence index plot shown in Figure 1 summarizes the outcomes of all 14 tasks for all sample members. The y-axis represents the IP members and the x-axis the additional tasks. Each horizontal line represents the outcomes for the temporal sequence of tasks for a sample member. For example, individuals in the bottom rows were ineligible for the first 13 tasks and participated in the 14th task. The first striking result is that the sequences are dominated by tasks respondents were ineligible for, illustrating the churn of sample members flowing in and out of the panel. One of the more frequent patterns is individuals initially participating in the first one to three tasks, followed by non-participation and then becoming ineligible (mostly due to drop out from the panel). Another frequent participation pattern is respondents alternating, participating in a sub-set of tasks, in no discernible systematic way.

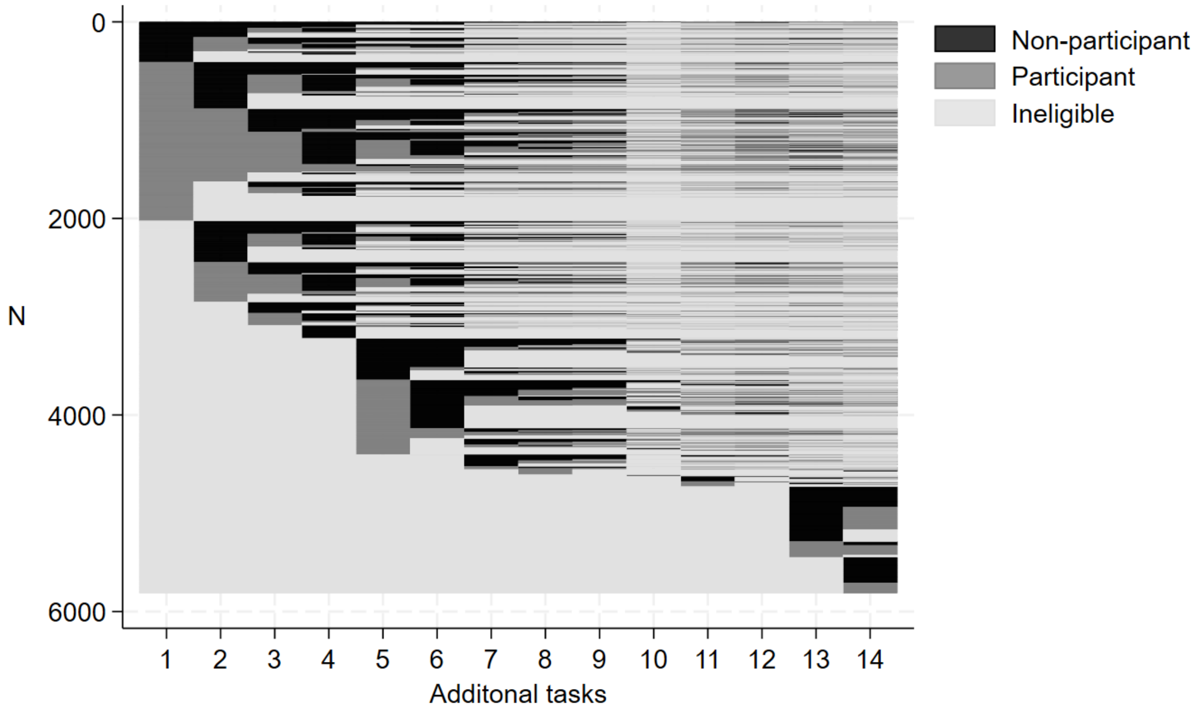


Figure 1. Sequence analysis of participation patterns in the Innovation Panel (IP6 – IP16) multi-modal set of tasks.

Next, we focus on participation in tasks sample members were eligible for. Figure 2 shows the mean participation rates and associated 95% confidence intervals, by number of tasks the sample member was invited to. Sample members participated in just under half of the tasks they were invited to (see Appendix H). The participation rates increase slightly with the number of tasks eligible for.

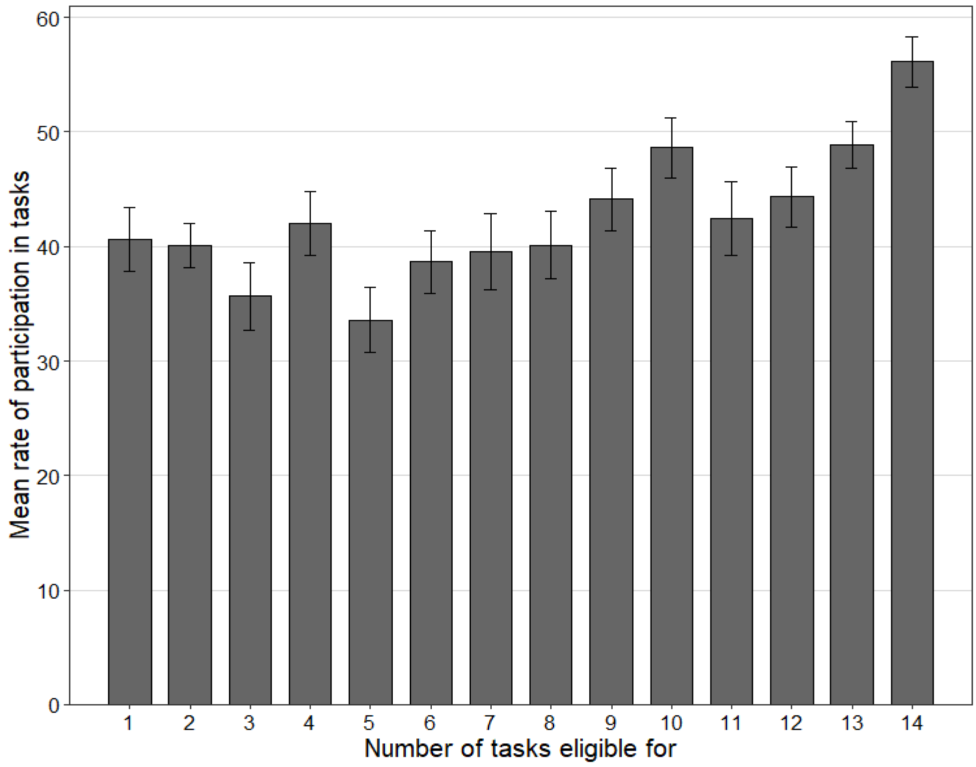


Figure 2. Mean participation rates in 14 additional tasks in the Innovation Panel (IP6 – IP16), by eligibility. Error bars represent confidence intervals.

Figure 3 shows patterns of participation by task type and number of tasks for which the sample member was eligible. For example, the last bar shows that of the sample members eligible for all 14 tasks, 69.5% participated in all three types of tasks, and 24.8% participated in within interview and outside interview tasks only. This pattern appears consistent across the different eligibility groups. Overall, the results suggest that as sample members were invited to more tasks, they were less likely to participate in none or only one type of task, and more likely to participate in all three types of tasks. The analyses by task topic and task incentive show similar results: as the number of tasks eligible for increased, the percentage of those participating in all topics (health, finance, events) and both incentivised and non-incentivised tasks increased (see Appendix I and J). Focusing on the panel members eligible for all 14 tasks, nearly 70% participated in all three types of tasks, topics, and incentivised and non-incentivised tasks.

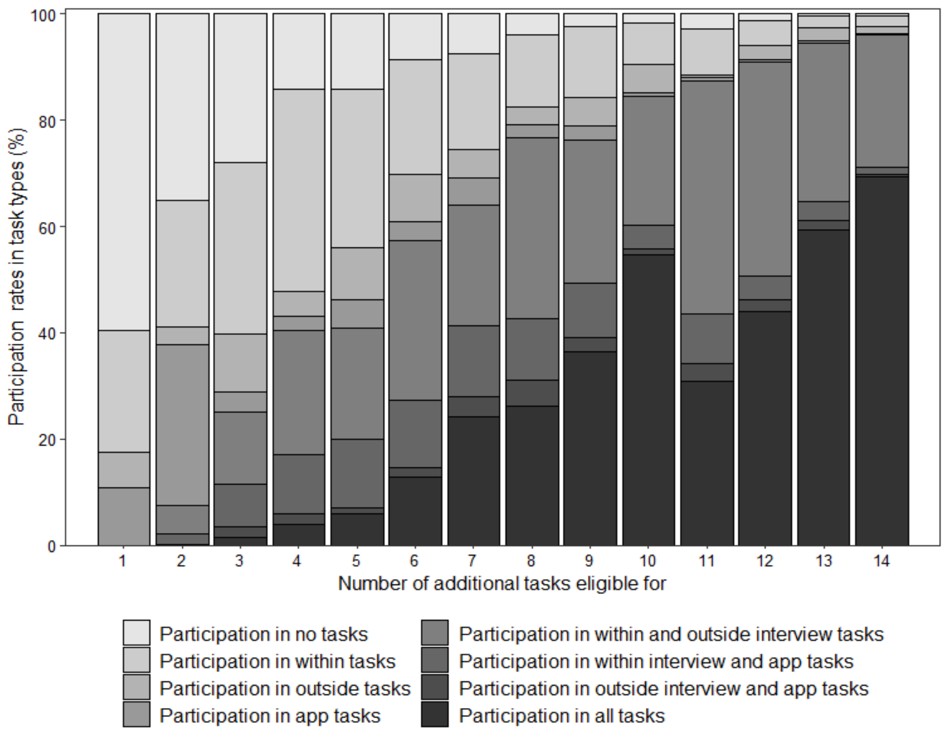


Figure 3. Participation rates in types of tasks in the Innovation Panel (IP6 –IP16), by eligibility.

Focusing on the 347 sample members eligible for all tasks, there are no prevalent patterns of which tasks they participated in (Appendix K). The most frequent pattern was respondents participating in all tasks except three of the app tasks, but this represented only 1.7% of respondents eligible for all tasks.

In sum, there was a high rate of churn within the sample, with individuals flowing in and out of the panel, initially participating followed by non-participation, then dropout. Regardless of how many tasks sample members were invited to, they participated in just under half the additional tasks they were eligible for, with no systematic patterns of participation in the set of multi-modal tasks.

5.2 Do respondents become less likely to participate in additional tasks the more tasks they are asked to do?

Figure 4 shows the participation rates in the additional tasks dependent on how many tasks the sample members were invited to previously. Participation rates fluctuated, indicating varying levels of engagement in the additional tasks as the number of previous invitations increased.

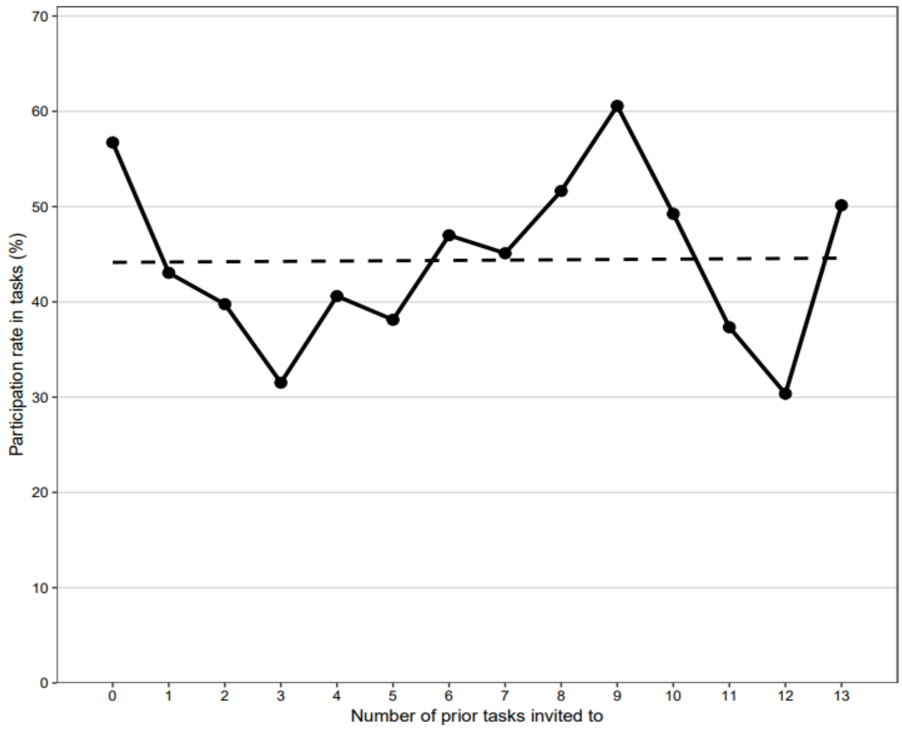


Figure 4. Participation of additional tasks in the Innovation Panel (IP6 – IP16), by number of previous tasks invited to.

Table 2 documents the results from the models predicting participation (see Appendix L for the full table). An initial model regressing the probability of participation on the number of prior tasks invited to produced a coefficient of $\beta = -0.003$ ($p = 0.003$). We then tested adding various controls. Controlling for the current task resulted in an improved model fit, as evidenced by the reduced AIC and BIC values in Appendix M. Controlling for sample origin did not improve model fit, however, the control was still included due to the presence of potential selection bias.

According to the unweighted model, the probability of participating in a task decreased slightly by 0.1 percentage points with each additional task the respondent was invited to previously. However, this coefficient was non-significant ($p = 0.729$), suggesting there is no association between number of prior tasks invited to and participation in the next task. Similarly, the weighted and item non-response models were non-significant, with $\beta = -0.001$.

The β coefficient estimated from the fixed-effects and instrumental models were, however, significant. The fixed-effects model found a one percentage increase in the probability of participation with each additional previous task invitation. However, the instrumental regression estimated a significant negative association between participation in a task and the number of previous tasks invited to. That is, the probability of participating in a task decreased by four percentage points with each additional task the respondent was invited to previously.

Comparing the different methods, using non-response weights can only account for differences in terms of variables measured in the survey. That is, this method is unlikely to account for all relevant aspects of selection in the continuing sample. The fixed effects method can only account for unobservable differences between sample members that are fixed over time and is therefore also unlikely to account for all differences. In addition, the model drops all cases where the individual participated in all the tasks or did not participate in any of the tasks, due to no within variation (Kennedy, 2008). This may explain the positive effect estimated from the fixed effects model. Using item non-response as an indicator of cooperativeness provides a time-varying measure of cooperativeness, but also does not account for all aspects of selection.

In contrast, the instrumental variable approach, estimated with a strong instrument, accounts for all potential differences between those remaining in the panel and those dropping out – whether in observable or unobservable characteristics. This is therefore our best estimate of the effect of the number of prior tasks invited to on participation (Angrist & Pischke, 2009; Wing, 2019). Our findings support this, with only the instrumental regression showing a negative relationship, suggesting the other methods did not full account for selectivity.

Table 2

Linear probability models of number of prior tasks invited to on participation in a task (IP6 – IP16).

	Unweighted	Weighted	Fixed-effects	Item non-response control	Instrumental regression
Number of prior tasks β	-0.001	-0.001	0.010	-0.001	-0.040
Standard error	0.002	0.003	0.004	0.002	0.003
P-value	0.729	0.631	0.010	0.540	0.000
Controls:					
Task	Yes	Yes	Yes	Yes	Yes
Sample Origin	Yes	Yes	No	Yes	No
Item non-response	Yes	Yes	No	Yes	Yes
Mode	Yes	Yes	No	Yes	Yes
Task experiments	Yes	Yes	No	Yes	Yes
AIC	37,716.75	-	28,688.17	35,436.62	-
N	31,728	29,949	27,758	29,949	31,728

Coefficients represent the effect of the number of previous task invitations on participation in the next task. Unweighted = Standard error adjusted for clustering in individuals (pidp). Weighted = Innovation Panel sample members who completed the annual interview. Non-response weights. Standard error adjusted for clustering and stratification. Fixed-effects

model = 3,970 cases dropped due to no within variation. Task experiment controls excluded due to collinearity. Non-response indicator model = Item non-response in the corresponding annual interview of sample members who completed the annual interview.
Instrumental regression = Sample origin used as an instrument for the number of prior tasks invited to.

5.3 Do respondents become less likely to participate in the annual interview the more additional tasks they are asked to participate in?

Figure 5 shows the participation rates for the annual interview by number of prior tasks the respondent was invited to. The positive relationship illustrates the potential presence of self-selection bias: sample members who are invited to more tasks are also more likely to complete the annual interview.

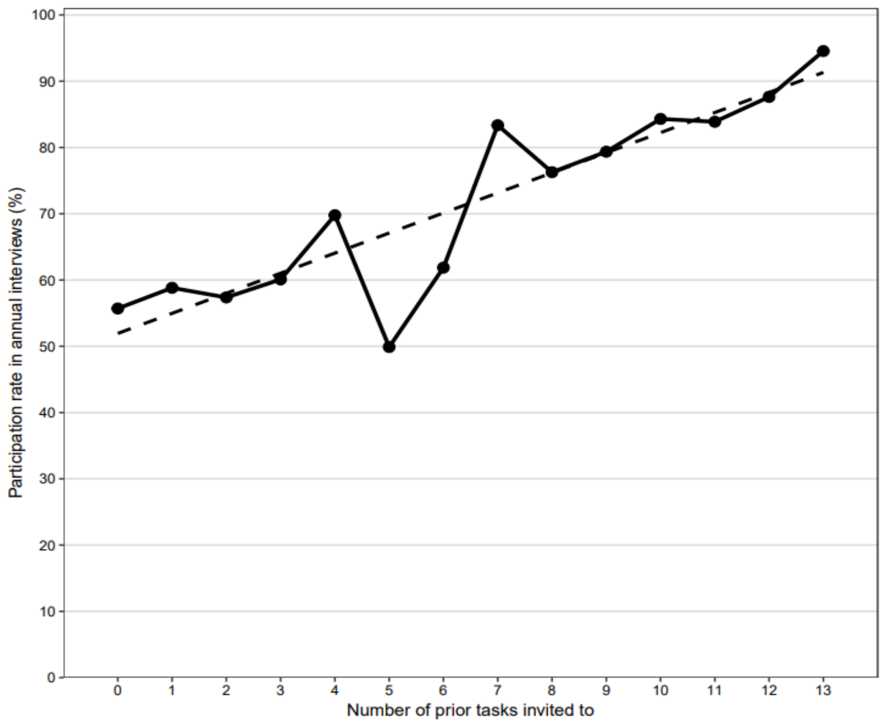


Figure 5. Participation in the Innovation Panel annual interviews (IP1 – IP16), by number of previous tasks invited to.

We estimated several alternative regressions to predict participation in the annual interviews based on the number of previous tasks sample members were invited to (Table 3;

for the full table see Appendix N). Adding controls for interview year, sample origin and the interview experiments produced a better model fit (see Appendix O).

All regressions yielded Wald tests for the β coefficient with p-values less than 0.001, except for the instrumental regression, which had a p-value of 0.037. This indicates that the number of previous task invitations significantly influenced participation in the annual interviews. Focusing on each model, the unweighted regression demonstrated that the probability of participating in an annual interview increased by 8.6 percentage points with each additional prior task invitation. Similarly, the weighted model estimated that the likelihood of participating in an annual interview increased by 6.9 percentage points with each additional previous task request. The fixed-effects model estimated a smaller positive effect of 1.5 percentage points. However, the positive association did not hold in the instrumental model, which found a significant, slightly negative, association. That is, the probability of participating in the annual interview decreased by 0.5 percentage points with each additional task invitation.

As with the analyses for RQ2, the positive association between number of prior tasks invited to and probability of participation in the annual interview is likely due to the different methods not fully accounting for selection bias. Our best estimate is again the instrumental variable estimate.

Overall, therefore, the raw data show a positive association between the number of prior tasks invited to and the probability of participating in annual interviews. Controlling for selection bias due to drop out of less cooperative respondents from the panel, however, suggests that each additional task invite in fact has a negative, albeit small, effect on retention in the panel.

Table 3

Linear probability models of number of prior tasks invited to on participation in the annual interview (IP1 – IP16).

	Unweighted	Weighted	Fixed-effects	Instrumental regression
Number of prior tasks β	0.086	0.069	0.015	-0.005
Standard error	0.001	0.002	0.002	0.003
P-value	0.000	0.000	0.000	0.037
Controls				
Interview year	Yes	Yes	Yes	Yes
Sample origin	Yes	Yes	No	No
Incentives	Yes	Yes	Yes	Yes
Mode	Yes	Yes	Yes	Yes
Individual mailings experiment	Yes	Yes	Yes	Yes
Household mailings experiment	Yes	Yes	Yes	Yes
Compression experiment	Yes	Yes	Yes	Yes
Advance letter wording experiment	Yes	Yes	Yes	Yes
Wave 2 mode experiment	Yes	Yes	No	Yes
AIC	67,493.76	-	38,397.72	-
N	61,178	44,534	41,154	61,178

Coefficients of the number of previous task invitations on participation in the annual interview. All models control included dummy indicators for what the current interview year was. 959 cases dropped due to no incentive information. Unweighted = Standard error adjusted for clustering in individuals (pidp). Weighted model = Standard error adjusted for clustering and stratification using enumerated weights. 16,644 cases dropped due to no enumeration weights. Fixed-effects model = 20,024 cases dropped due to no within variation. Instrumental model = Sample origin used as an instrument for the number of prior tasks invited to.

6. Discussion

As respondents are increasingly asked to do more than answer survey questions, it is important to examine what effects additional data collection tasks have on people's willingness to continue to contribute to a survey. We used 10 years of annual data from the *Understanding Society* Innovation Panel, a probability sample of households in Great Britain. This included 14 additional data collection tasks: consents to data linkage, several mobile app studies, monthly mini-surveys, time diaries, bio-measure and finger measurements. We used these data to examine the patterns of participation across different additional data collection tasks in a panel study, and to estimate the effect of previous task requests on participation in later tasks and annual interviews.

Examining the patterns of participation across tasks (RQ1) we found high churn in the sample, with individuals tending to participate in just under half of the tasks they were invited to. This suggests that IP members may have alternated participating in the tasks. However, this does not hold as there were no clear alternating patterns of participating for most sample members. The only consistent pattern was that regardless of how many tasks respondents had been invited to, around a quarter only participated in the within interview and outside interview with supplies provided tasks. This can, however, be explained by the generally low uptake of app tasks.

Examining whether respondents become less likely to participate in additional tasks the more tasks they are asked to do (RQ2), initial analyses found either no or a positive effect. However, the initial analyses did not fully control for potential selection bias: respondents who had been invited to more tasks were likely more cooperative than those invited to fewer tasks, since they had been in the panel for longer. The instrumental variable regression, our best estimate as described in section 5.2, found a negative relationship. This suggests that each additional task request reduces the likelihood of participation in the next task by 4 percentage points.

Examining whether respondents become less likely to participate in the annual interview the more additional tasks they are asked to participate in (RQ3) produced similar estimates. The instrumental variable model estimated that each additional task invitation reduces the probability of participating in the next annual interview by 0.5 percentage points.

Other panel studies have found similar results in the initial phase of multi-modal datasets. That is, participation in a survey wave slightly decreases, by no more than four percentage points, when individuals were invited to participate in one extra task (Eisnecker & Kroh, 2016; A Jäckle, J Burton, M.P Couper, Vine. J, et al., 2023; Trappmann et al., 2023). However, the current research goes beyond the existing literature to include more than one additional task and finding that these initial results do not hold when the number of previous task requests increases. Although our results for a single task are similar to those in the previous literature, we find cumulative effects: 14 additional tasks reduce participation in the next task by an estimated 52 percentage points and participation the next annual interview by 7 percentage points. That is, as the number of additional tasks increases over time, the effects on participation and panel retention are no longer small.

These findings have implications for panels that were not previously known. Inviting respondents to participate in additional tasks might be detrimental to survey panels. Whilst one or two additional tasks will not cause concerning levels of non-response or attrition, when tasks continue to be added, the panels response rates may significantly reduce. This leaves a trade-off between collecting more (reliable) data from additional tasks and increasing non-response and attrition.

Ideally, there would have been a control group enabling a comparison of participation between those invited to the additional tasks and those not invited. We did the next best in order to identify causal effects of task invitations on participation and attrition. We exploited exogenous variation in the number of tasks respondents were invited to, provided by the addition of refreshment samples added at regular intervals. This meant that sub-sets of the sample had been invited to different numbers of tasks, not due to potentially selective attrition from the panel, but due to random selection. The sample origin was therefore a strong instrument for instrumental variable regressions, estimating the causal effects of prior task invitations on participation and attrition.

There are further questions that we did not investigate, such as participation biases in the multi-modal set of tasks. From previous research, it can be inferred that selection biases might be present as they are present in the tasks individually (Abraham et al., 2006; Jäckle et al., 2019; Máté et al., 2023; Te Braak et al., 2020). Additionally, we did not determine which types of tasks are more likely to cause dropout. That is, which tasks are safe to include in a panel, and which tasks are more risky due to their negative effects on panel retention?

Future research, therefore, is needed to identify selection biases in sets of multi-modal tasks. Specifically, how the extent and the nature of selection biases compare between the additional tasks. Furthermore, if data users wanted to combine data collected in different ways, would their resulting analysis sample be more or less selective than if they used data from a single additional task? Are those who are more likely to drop out due to additional tasks the same types of respondents who are already under-represented in the panel?

Our results therefore suggest that while new methods and technologies offer exciting opportunities to collect novel or more accurate information about sample members than survey questions can, there are important trade-offs to consider for the quality of the resulting data.

Bibliography

Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Nonresponse in the American time use survey: Who is missing from the data and how much does it matter? *International Journal of Public Opinion Quarterly*, *70*(5), 676-703.

Al Baghal, T., Burton, J., Crossley, T. F., Benzeval, M., & Kumari, M. (2025). How Different Mixed-Mode Data Collection Approaches Impact Response Rates and Provision of Biomeasure Samples. *Public Opinion Quarterly*, *89*(2), 245-269. https://doi.org/10.1093/poq/nfaf022

Allum, N., Auspurg, K., Blake, M., Booker, C. L., Crossley, T. F., d'Ardenne, J., Fairbrother, M., Iacovou, M., Jäckle, A., Kaminska, O., Lynn, P., Nicoletti, C., Oldfield, Z., Pudney, S., Schnettler, S., Uhrig, N. S. C., & Winter, J. (2014). 'Understanding Society Innovation Panel Wave 6: results from methodological experiments.' Understanding Society Working Paper 2023 – 07. Colchester: University of Essex. Available at https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2014-04

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* . Princeton university press.

Benzeval, M., Aguirre, E., & Kumari, M. (2023). Understanding Society: health, biomarker and genetic data. *Fiscal Studies*, *44*(4), 399-415.

Boyle, J., Berman, L., Dayton, J., Iachan, R., Jans, M., & ZuWallack, R. (2021). Physical measures and biomarker collection in health surveys: Propensity to participate. *Research in Social and Administrative Pharmacy*, *17*(5), 921-929.

Burton, J., Couper, M., & Jäckle, A. (2024). 'The effects of placement and order on consent to data linkage in a web survey', *Journal of Survey Statistics and Methodology,* smae004. Available at: https://doi.org/10.1093/jssam/smae004

Burton, J., Jäckle, A., & Couper, M. (2024). '10. The Effect of Incentives on the Willingness to Participate in a Gamified App Study'. In Understanding Society Innovation Panel Wave 16: Results from Methodological Experiments and New Data, by Jim Vine et al. Understanding Society Working Paper Series 2024–11. Colchester, Essex: Institute for Social and Economic Research, University of Essex.

Cooke, M., Nielsen, A., & Strong, C. (2003). 'The Use of SMS as a Research Tool', *Proceedings of the Fourth ASC International Conference*, September. pp. 267-276.

Couper, M. P. (2017). New developments in survey data collection. *Annual review of sociology*, *43*(1), 121-145.

Di Matteo, D., Fine, A., Fotinos, K., Rose, J., & Katzman, M. (2018). 'Patient willingness to consent to mobile phone data collection for mental health apps: structured questionnaire', *JMIR Mental Health*, 5(3). Available at: 10.2196/mental.9539

Eggleston, J. (2024). Frequent Survey Requests and Declining Response Rates: Evidence from the 2020 Census and Household Surveys. *Journal of survey statistics and methodology*, *12*(5), 1138-1156. https://doi.org/10.1093/jssam/smae022

Eisnecker, P. S., & Kroh, M. (2016). The informed consent to record linkage in panel studies: Optimal starting wave, consent refusals, and subsequent panel attrition. *Public Opinion Quarterly*, *81*(1), 131-143.

Fitzsimons, E., Moulton, V., Hughes, D., Neaves, S., Ho, K., Hemani, G., Timpson, N., Calderwood, L., Gilbert, E., & Ring, S. (2020). 'Collection of DNA samples and genetic data at scale in the UK Millennium Cohort Study'. CLS working paper number 2020/7. UCL Institute of Education, University College London. Avaliable at: https://cls.ucl.ac.uk/wp-content/uploads/2020/08/CLS-working-paper-2020-7-Collection-of-DNA-samples-and-genetic-data-at-scale-in-the-UK-Millennium-Cohort-Study.pdf

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

Gerber, A. S., & Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. W.W. Norton & Company.

Groves, R. M., & Couper, M. P. (2012). *Nonresponse in Household Interview Surveys*. John Wiley & Sons.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (Vol. 561). John Wiley & Sons.

Hoe, N. D., & Grunwald, H. E. (2015). 'The role of automated SMS text messaging in survey research', *Survey Practice*, 8(6). Available at: https://doi.org/10.29115/SP-2015-0027

Ingen, E. V., Stoop, I., & Breedveld, K. (2009). Nonresponse in the Dutch time use survey: Strategies for response enhancement and bias reduction. *Field Methods*, *21*(1), 69-90.

Institute for Social and Economic Research. (2023). Understanding Society – The UK Household Longitudinal Study, Innovation Panel, Waves 1-15, User Manual. Colchester: University of Essex. Avaliable at: https://www.understandingsociety.ac.uk/documentation/innovation-panel/user-guide/

Institute for Social and Economic Research. (2024). Understanding Society – The UK Household Longitudinal Study, Innovation Panel, Waves 1-16, User Manual. Colchester: University of Essex. Avaliable at: https://www.understandingsociety.ac.uk/documentation/innovation-panel/user-guide/

Jäckle, A., Burton, J., & Couper, M. P. (2023). Understanding Society: minimising selection biases in data collection using mobile apps. *Fiscal Studies*, *44*(4), 361-376.

Jäckle, A., Burton, J., Couper, M. P., Crossley, T. F., & Walzenbach, S. (2024). Survey consent to administrative data linkage: five experiments on wording and format. *Journal of survey statistics and methodology*, *12*(5), 1174-1199.

Jäckle, A., Burton, J., Couper, M. P., J, V., & Horn, D. (2023). 'Trialling event-triggered data collection in the Understanding Society Innovation Panel: response and measurement quality.' Understanding Society Working Paper 2023-09. Colchester: University of

Essex. Available at https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2023-09

Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). 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., Burton, J., Couper, M. P., & Perelli-Harris, B. (2023). 'Participation of household panel members in daily burst measurement using a mobile app: effects of position of the invitation, bonus incentives, and number of daily questions'. Understanding Society Working Paper 2023-01. Colchester: University of Essex. Available at https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2023-01.pdf

Jäckle, A., Wenz, A., Burton, J., & Couper, M. P. (2022). Increasing participation in a mobile app study: the effects of a sequential mixed-mode design and in-interview invitation. *Journal of survey statistics and methodology*, *10*(4), 898-922.

Jenkins, S. P., Cappellari, L., Lynn, P., Jäckle, A., & Sala, E. (2006). Patterns of Consent: Evidence from a General Household Survey. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *169*(4), 701-722. https://doi.org/10.1111/j.1467-985X.2006.00417.x

Kearney, P. M., Cronin, H., O'Regan, C., Kamiya, Y., Whelan, B. J., & Kenny, R. A. (2011). Comparison of centre and home-based health assessments: early experience from the Irish Longitudinal Study on Ageing (TILDA). *Age and ageing*, *40*(1), 85-90.

Kennedy, P. (2008). *A Guide to Econometrics*. John Wiley & Sons.

Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., Trappmann, M., & Eckman, S. (2022). Non-Participation in Smartphone Data Collection Using Research Apps. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*(2), 225-245. https://doi.org/10.1111/rssa.12827

Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly*, *83*(S1), 210-235. https://doi.org/10.1093/poq/nfz007

Link, M. W., Murphy, J., Schober, M. F., Buskirk, T. D., Hunter Childs, J., & Langer Tesfaye, C. (2014). Mobile Technologies for Conducting, Augmenting and Potentially Replacing Surveys: Executive Summary of the AAPOR Task Force on Emerging Technologies in Public Opinion Research. *Public Opinion Quarterly*, *78*(4), 779-787. https://doi.org/10.1093/poq/nfu054

Lynn, P. (2014). Longer interviews may not affect subsequent survey participation propensity. *Public Opinion Quarterly*, *78*(2), 500-509.

Máté, Á., Rakovics, Z., Rudas, S., Wallis, L., Ságvári, B., Huszár, Á., & Koltai, J. (2023). Willingness of participation in an application-based digital data collection among different social groups and smartphone user clusters. *Sensors*, *23*(9), 4571.

Morris, S., Humphrey, A., Cabrera Alvarez, P., & D'Lima, O. (2016). *The UK Time Diary Study 2014-2015.* NatCen Technical Report. Available at: https://doc.ukdataservice.ac.uk/doc/8128/mrdoc/pdf/8128_natcen_reports.pdf

Mostafa, T. (2016). Variation within households in consent to link survey data to administrative records: evidence from the UK Millennium Cohort Study. *International Journal of Social Research Methodology*, *19*(3), 355-375. https://doi.org/10.1080/13645579.2015.1019264

Mostafa, T., & Wiggins, R. D. (2018). What influences respondents to behave consistently when asked to consent to health record linkage on repeat occasions? *International Journal of Social Research Methodology*, *21*(1), 119-134. https://doi.org/10.1080/13645579.2017.1332848

Office for National Statistics. (2003). *UK 2000 Time Use Survey - Summary of main changes for the delivery to Essex Data Archive, July 2003.* Available at: https://doc.ukdataservice.ac.uk/doc/4504/mrdoc/pdf/4504userguide1.pdf

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358-369.

Revilla, M., Couper, M. P., & Ochoa, C. (2019). Willingness of online panelists to perform additional tasks. *Methods, data, analyses*, *13*(2), 223-252.

Revilla, M., Paura, E., & Ochoa, C. (2021). Use of a research app in an online opt-in panel: The Netquest case. *Methodological Innovations*, *14*(1), 2059799120985373.

Robinson, J. P. (1998). Activity patterns of time-diary dropouts. *Loisir et société/Society and Leisure*, *21*(2), 551-554.

Sakshaug, J., Tutz, V., & Kreuter, F. (2013). Placement, wording, and interviewers: Identifying correlates of consent to link survey and administrative data. *Survey research methods*, *7*(2), 133-144.

Sakshaug, J. W., Couper, M. P., & Ofstedal, M. B. (2010). Characteristics of physical measurement consent in a population-based survey of older adults. *Medical care*, *48*(1), 64-71.

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

Sakshaug, J. W., Ofstedal, M. B., Guyer, H., & Beebe, T. J. (2014). 'The collection of biospecimens in health surveys', in T.P. Johnson (ed.). *Health Survey Methods.* Hoboken, New Jersey: John Wiley & Sons, pp. 383-419.

Sharp, L. M., & Frankel, J. (1983). Respondent burden: A test of some common assumptions. *Public Opinion Quarterly*, *47*(1), 36-53.

Struminskaya, B., Lugtig, P., Keusch, F., & Höhne, J. K. (2020). Augmenting Surveys with Data from Sensors and Apps: Opportunities and Challenges. *Social Science Computer Review*, 1-13.

Te Braak, P., Minnen, J., & Glorieux, I. (2020). The representativeness of online time use surveys. Effects of individual time use patterns and survey design on the timing of survey dropout. *Journal of Official Statistics*, *36*(4), 887-906.

The American Association for Public Opinion Research. (2023). . *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 10th edition. AAPOR. Avaliable at: https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf

Torous, J., Lipschitz, J., Ng, M., & Firth, J. (2020). Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *Journal of affective disorders*, *263*, 413-419.

Trappmann, M., Haas, G.-C., Malich, S., Keusch, F., Bähr, S., Kreuter, F., & Schwarz, S. (2023). Augmenting survey data with digital trace data: Is there a threat to panel retention? *Journal of survey statistics and methodology*, *11*(3), 541-552.

University of Essex, Institute for Social and Economic Research. (2024). Understanding Society: Innovation Panel, Waves 1-16, 2008-2022, [data collection], UK Data Service, 12th Edition, Accessed 14 November 2024. SN: 6849, DOI: http://doi.org/10.5255/UKDA-SN-6849-15

University of Essex, Institute Social Economic Research. (2022a). Understanding Society: Spending Study 1, 2016-2017, [data collection], UK Data Service, Accessed 10 January 2024. SN: 8749, DOI: http://doi.org/10.5255/UKDA-SN-8749-1

University of Essex, Institute Social Economic Research. (2022b). Understanding Society: Spending Study 2, 2018-2019, [data collection], UK Data Service, Accessed 10 January 2024. SN: 8909, DOI: http://doi.org/10.5255/UKDA-SN-8909-1

Verbeek, M. (2000). *A Guide to Modern Econometrics*. John Wiley & Sons.

Vine, J., Baghal, T. A., Burton, J., Couper, M. P., Crossley, T. F., Fisher, P., Gush, K., Hristova, D. H., Jäckle, A., Jessop, C., Kumari, M., Levell, P., Low, S. L., Parbury, S., Serodio, P., Sloan, L., & Walzenbach, S. (2023). 'Understanding Society Innovation Panel wave 15: results from methodological experiments.' Understanding Society Working Paper 2023-10. Colchester: University of Essex. Available at https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2023-10

Vine, J., Jäckle, A., Burton, J., & Couper, M. P. (2025). 'Getting consent to survey in new ways: evidence from experiments about questions by text message in the Understanding Society Innovation Panel', Understanding Society Working Paper Series No. 2025-14, Colchester: University of Essex. Avaliable at: https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2025-14/

Vingeliene, S., Hiyoshi, A., Lentjes, M., Fall, K., & Montgomery, S. (2019). Longitudinal analysis of loneliness and inflammation at older ages: English longitudinal study of ageing. *Psychoneuroendocrinology*, *110*, 104421. https://doi.org/https://doi.org/10.1016/j.psyneuen.2019.104421

Wing, C. (2019). What Can Instrumental Variables Tell Us About Nonresponse In Household Surveys and Political Polls? *Political Analysis*, *27*(3), 320-338.

Yan, T., & Williams, D. (2022). Response burden–Review and conceptual framework. *Journal of Official Statistics*, *38*(4), 939-961.

Appendix

Appendix A

Examples of additional task participation rates in existing literature.

Additional task	Study/sample	Year	Survey type	Topic/specific task	Participation/ willingness rates (%)	Reference
Bio-measures						
	MFour mobile panel	2018	Longitudinal study	Blood pressure, finger-stick blood draw, blood pressure machine	47 - 75	Boyle et al. (2021)
	Millennium Cohort Study	2022	Longitudinal study	Saliva	81	Fitzsimons et al. (2020)
	Irish Longitudinal Study on Ageing	2011	Longitudinal study	Physical assessment	61	Kearney et al. (2011)
	English Longitudinal Study of Ageing	2012/ 2013	Longitudinal study	Anthropometric measurements, blood pressure, whole blood, saliva	84	Vingeliene et al. (2019)
	Health and Retirement Study	2006	Panel survey	Saliva, blood spots	Saliva = 84, blood spot = 83	Sakshaug et al. (2010)
Data linkage consent						
	Understanding Society	2024	Longitudinal study	NHS health records	71	Burton, Couper, et al. (2024)
	Millennium Cohort Study	2018	Longitudinal study	Children's health records	76	Mostafa and Wiggins (2018)

	Panel Study of Income Dynamics	2014	Longitudinal study	Medicare data	21	Freedman et al. (2014)
	Persons drawn from German federal databases used in the social security administration	2013	Cross-sectional study	Employment data	95	Sakshaug et al. (2013)
	Health and Retirement Study	2012	Longitudinal study	Earnings and benefit histories record	68	Sakshaug et al. (2012)
SMS consent						
	1,000 mobile users associated with the six zip codes that surround a Philadelphia park	2015	Cross-sectional study	Wissahickon Park SMS question consent	22	Hoe and Grunwald (2015)
	Understanding Society	2025	Longitudinal study	SMS question consent	74	Vine et al. (2025)
Time-use diaries						
	1975 Americans' Use of Time study	1975	Cross-sectional study	Completed four separate diaries for different days across the year	73	Robinson (1998)
	2004 American Time Use Survey	2004	Cross-sectional study	Time use information collected for one day with post telephone interview	55	Abraham et al. (2006)
	2014 UK Time Diary Study	2014	Cross-sectional household study	Completed two 24-hour time-use diaries	81	Morris et al. (2016)
	Dutch Time Use Survey	2005	Cross-sectional household survey	Time-use diary over two consecutive days	35	Ingen et al. (2009)

Finger Measurement Studies						
	Understanding Society	2013	Longitudinal study	Provided measurements of their second and fourth digits of both hands	73-74	Allum et al. (2014)
Mini-monthly surveys						
	Understanding Society	2020-2021	Longitudinal study	Complete monthly surveys on life events for one year	12 months completed = 59, 11 months completed = 21	A Jäckle, J Burton, M.P Couper, Vine. J, et al. (2023)
Mobile apps						
	Understanding Society	2016-2022	Longitudinal study	4 apps: Providing expenditure data, completing daily wellbeing questions, taking a photo of body outline and answering body and health questions	13 - 45	Annette Jäckle et al. (2023)
	Android device users from the German Panel Study Labour Market and Social Security	2018	Panel survey	Collected data over six months through short in-app surveys and five passive mobile data collection	Provided the different types of passively collected data = 12 - 13, provided all types of data at least once = 11	Keusch et al. (2022)
	Patients referred to a mental health clinic in Canada	2018	Patient survey	Mental health monitoring app installation willingness	84	Di Matteo et al. (2018)
	The Netquest app panellists	2018-2019	Online non-probability panel survey	Different app surveys	19 - 31	Revilla et al. (2021)
	Meta-analysis of RCTs of apps targeting depressive symptoms in adults	2019	Meta-analysis	A systematic review of RCTs of apps targeting depressive symptoms in adults	Around 50	Torous et al. (2020)

Appendix B

Understanding Society Innovation Panel additional tasks information.

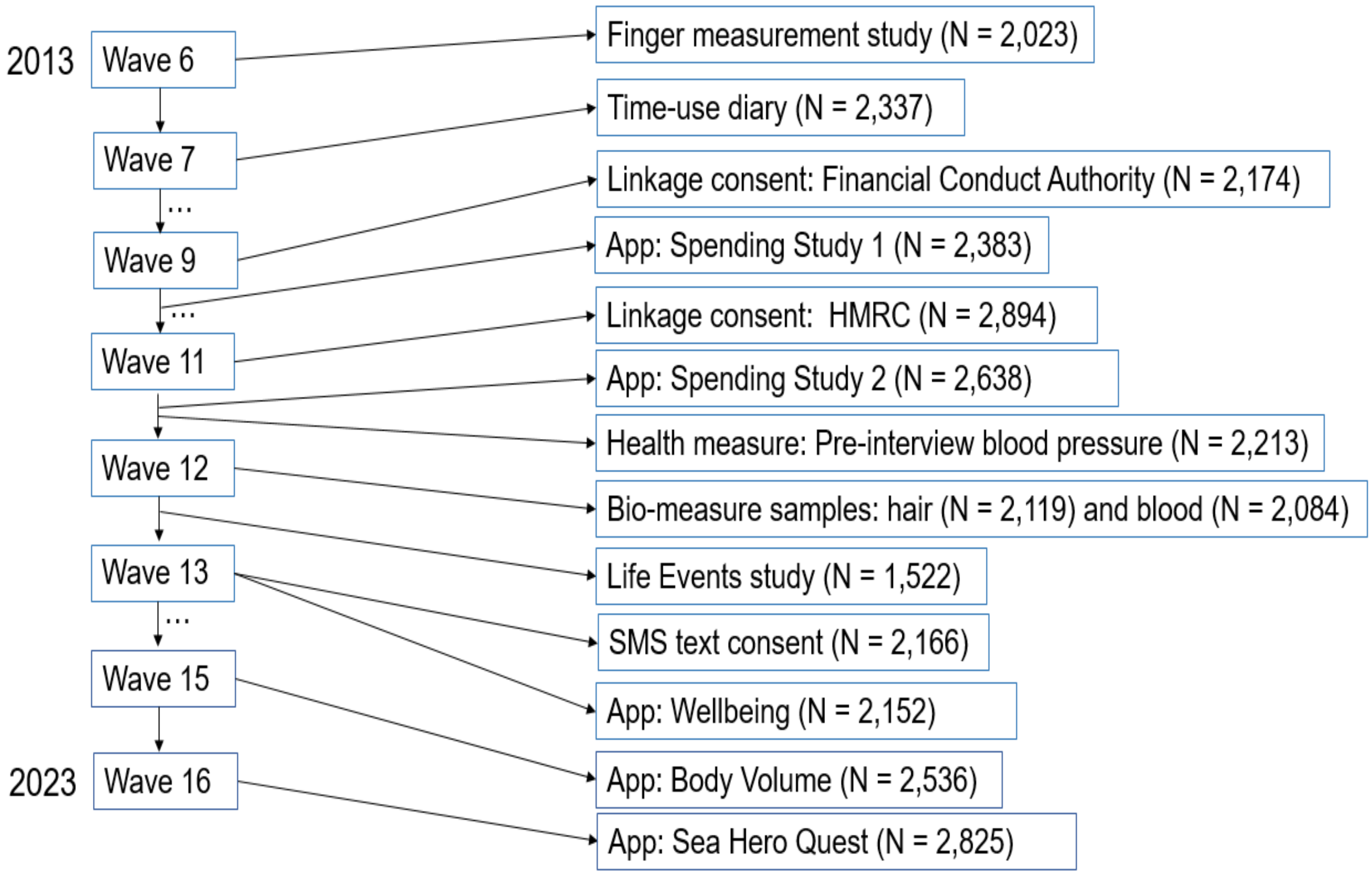
	Finger Measurement study	**Time Diary**	**FCA Consent**	**SS1**	**HMRC consent wave 11**	**SS2**	**Pre-interview blood pressure measure**	**Hair sample**	**Blood sample**	**Life Events study**	**Text consent**	**Wellbeing app**	**Body Volume app**	**Spatial Navigation Game**
Year	2013	2014	2016	2016	2018	2018	2019	2019	2019	2020 - 2021	2020	2020	2022	2023
Topic	Prenatal testosterone exposure and development	Time-use estimates	Financial Conduct Authority	Expenditure	HMRC records	Expenditure	Health measurements	Health biomarkers	Health biomarkers	Life events	SMS messaging mode	Wellbeing and relationships	Body measurements	Spatial cognition – Sea Hero Quest
Task	Measurement of finger-length ratio of second and fourth digit.	Complete two time use diaries indicating activities in 10-minute intervals from a pre-coded list (“light-touched”).	Consent to link FCA data to survey data	Upload receipts to app to report daily spending	Consent to link HMRC records to survey data	Complete diary on spending, direct debits and standing orders	Measure blood pressure in advance of the interview to provide blood measurements	Provide a strand of hair for health testing via a kit or nurse visit. Measurements of hair cortisol, cortisone, progesterone, and testosterone were measured.	Provide a dried blood spot sample for health testing via a kit or nurse visit Measurements of dried blood triglycerides, cholesterol, high-density lipoprotein cholesterol, and glycated haemoglobin were obtained.	Complete monthly surveys on events experienced in the month	Consent to be sent survey questions via SMS	Complete daily survey on wellbeing and relationships	Complete profile questions and take photos of oneself	Complete a navigation game with increasingly difficult levels.
Sample	IP6 respondents	IP 7 respondents	IP9 respondents	At least one household member who gave an IP9 interview	IP11 respondents	IP11 respondents	All individuals eligible for IP12	IP12 respondents who were not pregnant or breastfeeding and had hair	IP12 respondents who were not pregnant or breastfeedi	At least one household member who gave an IP11 interview	IP13 respondents who possessed a mobile phone	IP13 respondents who participated in IP12 or at least one	IP15 respondents who participated in IP14 or at least one	IP16 respondents

								longer than 2cm	ng, did not have a blood clot disorder, had not have a fit or mastectomy in past 5 and 1 year, and was not on renal dialysis or anti-coagulant	and in a household with regular internet use, had a known email or phone number and was allocated to sample rather than the control group who were not invited.		previous IP interview	previous IP interview	
Number of IP respondents invited	N = 2,023	N= 2,337	N = 2,174	N = 2,383	N = 2,894	N = 2,638	N = 2,213	N = 2,199	N = 2,084	N = 1,522	N = 2,166	N = 2,152	N = 2,536	N = 2,825
Invitation	Annual interview	Annual interview	Annual interview	Letter after annual interview	Annual interview	Experiment: letter vs. annual interview	Experiment: Letter before annual interview (including nearest pharmacy vs altruistic message vs control)	Annual interview	Annual interview	Email or SMS message	Annual interview	Annual interview	Annual interview	Annual interview
Incentives	No incentives	No incentives	No incentives	Experiment: £2 vs. £6 £1 for everyday app was used £10 for app used everyday	No incentives	50p for every day app used £1 for direct debit/standing order completion £3 debrief questionnaire completion £10 for app used everyday	Conditional £5 if blood pressure was measured.	£5 for biomarker completion	£5 for biomarker completion	Experiment: £1 for each monthly survey completed vs. £1 for completing the event question + £2 if they reported any events	No incentives	Experiment: £10 if app used every day vs. £2.50 on 4 random days if app was used vs. no additional incentive £1 every day app was used	Experiment: £5 if the app was used vs. £5 unconditional	Experiment: £10 vs. £30 if app was used.

Experiments	Annual interview incentives	Annual interview incentives	Annual interview incentives	Incentives	Annual interview incentives, wording of question, position in annual interview	Invitation to app, feedback (Lightspeed sample only).	Invitation to study, annual interview incentives	Annual interview incentives	Annual interview incentives	Incentives, timing of reminders	Position in annual interview, annual interview incentives, length of annual interview	Bonus incentives, length of app daily questionnaire, position in annual interview	Incentives, feedback	Incentives

Appendix C

Understanding Society Innovation Panel and additional tasks timeline, wave six to wave 16.



Appendix D

Table 4

Participation and eligibility status of the Innovation Panel additional tasks for all panellists who were eligible for at least one task (IP6 – IP16).

	Finger Measurement study		Time Diary		FCA Consent		Spending Study 1		HMRC consent (IP11)	
	N	%	N	%	N	%	N	%	N	%
Non-participant	408	7	1,050	18	935	16	1,842	32	1,242	21
Participant	1,615	28	1,287	22	1,239	21	270	5	1,652	28
Non-participant – wave NR/p/y	-	-	-	-	-	-	267	5	-	-
Participant – wave NR/p/y	-	-	5	0	-	-	4	0	-	-
Total eligible	2,023	35	2,337	40	2,174	37	2,383	41	2,894	50
Non-respondent – wave NR/p	219	4	752	13	419	7	107	2	1,027	18
Not invited – characteristics	-	-	-	-	-	-	-	-	-	-
Not invited – new household member	337	6	368	6	373	6	413	7	106	2
Not invited – refreshment sample	2,639	45	1,996	34	1,996	34	1,996	34	929	16
Not invited – experiment	-	-	-	-	-	-	-	-	-	-
Not invited – unknown	-	-	-	-	-	-	39	1	-	-
Ineligible – under 16	282	5	267	5	435	7	459	8	174	3
Ineligible – deceased	-	-	10	0	26	0	26	0	43	1
Ineligible – not in IP sample	312	5	77	1	17	0	17	0	10	0
Ineligible – removed from IP sample	-	-	-	-	372	6	372	6	629	11
Total ineligible	3,789	65	3,475	60	3,638	63	3,429	58	2,918	50
N	5,812	100	5,812	100	5,812	100	5,812	100	5,812	100

(Table 4 continued)

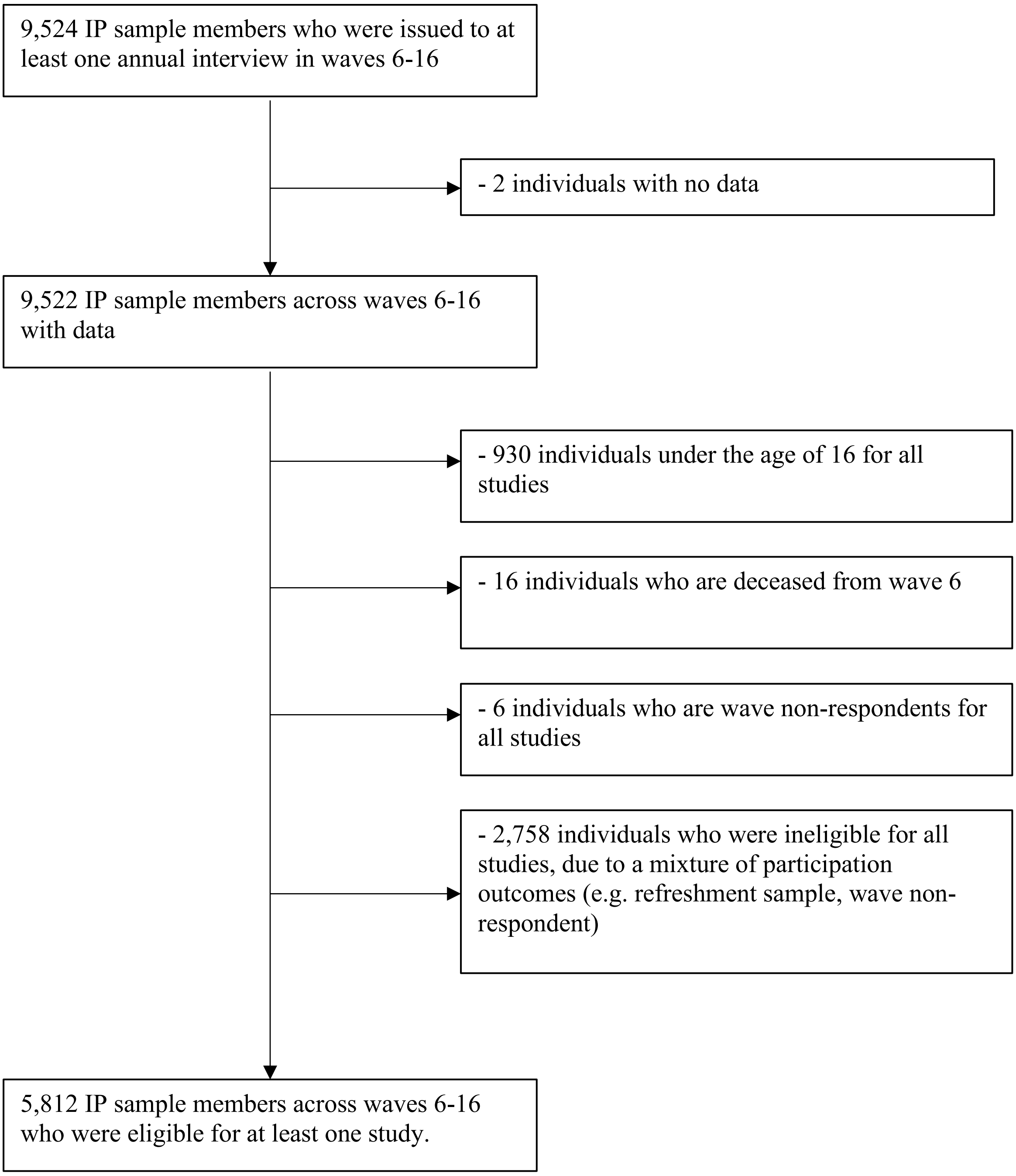
	Spending Study 2		Life events study		Pre-interview blood pressure		Hair sample		Blood sample	
	N	%	N	%	N	%	N	%	N	%
Non-participant	2,191	38	431	7	1,327	23	1,301	22	1,104	19
Participant	447	8	986	17	835	14	842	14	980	17
Non-participant – wave NR/p/y	-	-	69	1	51	1	-	-	-	-
Participant – wave NR/p/y	-	-	36	1	-	-	56	1	-	-
Total eligible	2,638	45	1,522	26	2,213	38	2,199	38	2,084	36
Non-respondent – wave NR/p	989	17	1,150	20	1,713	29	1,750	30	1,758	30
Not invited – characteristics	-	-	391	7	-	-	10	0	64	1
Not invited – new household member	105	2	110	2	59	1	79	1	79	1
Not invited – refreshment sample	1,239	21	929	16	929	16	929	16	929	16
Not invited – experiment	-	-	816	14	-	-	-	-		-
Not invited – unknown	-	-	17	0	-	-	-	-	-	-
Ineligible – under 16	159	3	169	3	134	2	81	1	134	2
Ineligible – deceased	43	1	64	1	64	1	64	1	64	1
Ineligible – not in IP sample	10	0	15	0	71	1	71	1	71	1
Ineligible – removed from IP sample	629	11	629	11	629	11	629	11	629	11
Total ineligible	3,174	55	4,290	74	3,599	62	3,613	62	3,728	64
N	5,812	100	5,812	100	5,812	100	5,812	100	5,812	100

(Table 4 continued)

	SMS text consent		Wellbeing app		Body Volume app		Spatial Navigation app	
	N	%	N	%	N	%	N	%
Non-participant	667	11	1,185	20	2,069	36	1,551	27
Participant	1,499	26	967	17	467	8	1,274	22
Non-participant – wave NR/p/y	-	-	-	-	-	-	-	-
Participant – wave NR/p/y	-	-	-	-	7	0	-	-
Total eligible	2,166	37	2,152	37	2,536	44	2,825	49
Non-respondent – wave NR/p	1,477	25	1,555	27	1,856	32	1,538	26
Not invited – characteristics	84	1	-	-	-	-	-	-
Not invited – new household member	42	1	61	1	16	0	1	0
Not invited – refreshment sample	929	16	929	16	-	-	-	-
Not invited – experiment	-	-	-	-	-	-	-	-
Not invited – unknown	2	0	1	0	-	-	-	-
Ineligible – under 16	92	2	92	2	52	1	13	0
Ineligible – deceased	87	2	89	2	107	2	115	2
Ineligible – not in IP sample	11	0	11	0	316	5	398	7
Ineligible – removed from IP sample	922	16	922	16	922	16	922	16
Total ineligible	3,646	63	3,660	63	3,276	56	2,987	51
N	5,812	100	5,812	100	5,812	100	5,812	100

Appendix E

Sample Flowchart for RQ1 and RQ2, including all Innovation panel members eligible for at least one additional task.



Appendix F

Table 5

Understanding Society Innovation Panel additional tasks experiments covariates.

Experiment	Description	Task	Treatment	Sample size	Unweighted regression F-test	Included in regressions?
SS1 incentives	Participants received a £2 or £6 incentive	SS1	£6	£2 = 1,160 £6 = 1,223	$F(2,5,708) = 112.15$, $p < 0.001$	Yes
HMRC consent question wording	Wording varied to explore the standard version used in the mainstage survey and a new easier to understand wording	HMRC	Easy version	Standard = 826 Easy = 770	$F(2,5,708) = 5.11$, $p = 0.006$	Yes
HMRC consent location	The location of the consent question was provided early to half the CAPI respondents and late to the other half	HMRC	Consent asked late	Early = 437 Late = 1,046	$F(2,5,708) = 1.86$, $p < 0.156$	No
SS2 invitation	Sample members were invited to SS2 in the wave 11 interview or in-between interviews with a postal letter	SS2	Invitation to download app made interwave	In-interview = 1,293 Interwave = 1,345	$F(2,5,708) = 5.09$, $p = 0.006$	Yes
Pre-interview blood pressure invitation	An advance letter experiment with one-third of the sample provided with information on their nearest pharmacy, one-third included an altruistic/pro-social appeal text, and the last third as the control group	Pre-interview blood pressure	Information on pharmacy Pro-social message	Information = 709 Pro-social = 726 Control = 778	$F(3,5,708) = 3.89$, $p = 0.009$	Yes

Bio-measures mode	Respondents were asked to provide bio-measures via a) a nurse b) an interviewer and self-kit c) themselves via a self-kit	Hair sample Blood sample	Nurse	Face-to-face = 671 Web = 796 Nurse = 732	Blood = $F(3,5{,}708) = 6.40$, $p = 0.003$ Hair = $F(3,5{,}708) = 12.10$, $p < 0.001$	Yes
Life Event Study incentives	Respondents received £1 with every completed monthly survey or £1 for completing the event question plus £2 if they reported any events	Life Events Study	£1 + £2	£1 = 742 £1+£2 = 780	$F(2,5{,}708) = 0.49$, $p = 0.615$	No
Life Events Study reminders	Sample members received a daily reminder or reminder every two days	Life Events Study	Reminders every 2 days	Reminders daily = 770 Reminders every 2 days = 752	$F(1,5{,}708) = 0.05$, $p = 0.821$	No
SMS text consent location	Text consent was asked either in the demographics module (early) or in the contact details module (late)	Text consent	Consent asked in Contact Details module	Early = 734 Late = 765	$F(2,5{,}708) = 377.43$, $p < 0.001$	Yes
Wellbeing app invitation location	Respondents were invited to the wellbeing app study early or late in the annual interview	Wellbeing app	Invitation to app late	Early = 1,098 Late = 1,054	$F(1,5{,}708) = 0.22$, $p = 0.639$	No
Wellbeing app incentives	Respondents received either £10 for all 14 days completed, £2.50 on four randomly selected days if they completed that day, or no additional incentive	Wellbeing app	£10 if all 14 days completed £2.50 on four randomly selected days if app survey completed on those days	No additional incentive = 740 £10 for all 14 days = 731 £2.50 on four randomly selected days = 681	$F(2,5{,}708) = 0.18$, $p = 0.833$	No
Wellbeing app length	Half the respondents received a two minute daily	Wellbeing app	10 minute daily app survey	2 minutes = 1,066	$F(2,5{,}708) = 21.96$, $p < 0.001$	Yes

	app survey and the other half received a 10 minute daily app survey			10 minutes = 1,086		
Body Volume app incentive	Respondents received a £5 conditional incentive or a 5 unconditional incentive based on participation	Body Volume app	Conditional £5 incentive	Unconditional £5 = 1,264 Conditional £5 = 1,272	$F(2,5{,}708) = 19.15$, $p < 0.001$	Yes
Body Volume app feedback	The feedback experiment meant participants received feedback on total body fat or feedback on visceral body fat or no feedback	Body Volume app	Feedback on total body fat Feedback on visceral body fat	No feedback = 822 Total body fat feedback = 863 Visceral body fat feedback = 851	$F(2,5{,}708) = 1.01$, $p = 0.364$	No
Spatial Navigation Game incentive	Respondents received a £10 conditional incentive for completing the game or a £30 conditional incentive for completing the game.	Spatial Navigation Game (Sea Hero Quest)	£30 conditional incentive	Inapplicable = 28 £10 = 2,142 £30 = 2,215	$F(2,5{,}708) = 8.43$, $p < 0.001$	Yes

Appendix G

Table 6

Understanding Society Innovation Panel annual interview experiments covariates.

Experiment	Description	IP wave	Treatment	Sample size	Unweighted regression F-test	Included in regressions?
Mixed mode experiment	Sample members were either in the mode face-to-face, issued to telephone and if one person could not be interviewed by telephone all remaining household members were transferred to CAPI or all household members were attempted to answer via telephone	Two	CATI move one, move all CATI try all	Face-to-face = 733 CATI move one, move all = 698 CATI try all = 713	$F(3,9{,}854)$ = 0.55.02, $p <$ 0.001	Yes
Self-completion	Respondents completed the self-completion section on paper or via CASI	Six	CASI	CASI = 1,420 Paper = 1,347	$F(1,9{,}854)$ = 0.12, p = 0.733	No
Targeted advanced letters	An experiment on the effect of using a tailored advance letter to the sample members demographic compared to a standard letter	Six	Targeted advance letter	Standard = 1,365 Tailored = 1,402	$F(1,9{,}854)$ = 0.10, p = 0.750	No
Frequency of inter-wave mailings	Allocation of households at random to receiving one (November) vs two to three (September, November, February) mailings between wave six and seven	Seven	Frequent mailings	Control = 1,344 Frequent mailings = 1,429	$F(1,9{,}854)$ = 1.81, p = 0.179	No

Targeted weekday invitation emails	Normal contact procedures were followed for half the mixed-mode sample, the other half received an email invitation on the day predicted to be more likely to lead to response. Individuals within these households who did not respond were sent a reminder email based on their preferred day	Nine	Sunday Monday Tuesday Wednesday Thursday Friday Saturday	Household: Sunday = 42 Monday = 63 Tuesday = 109 Wednesday = 172 Thursday = 109 Friday = 60 Saturday = 50 Control = 557 Individual: Sunday = 28 Monday = 63 Tuesday = 83 Wednesday = 119 Thursday = 77 Friday = 59 Saturday = 39 Control = 414	Household = $F(9,9{,}854) = 11.15$, $p < 0.001$ Individuals = $F(9,9{,}854) = 13.07$, $p < 0.001$	Yes
Advance letter wording	Household were randomly assigned to receive a positive outcome wording advance letter appealing to altruism or a negative outcome eroding letter.	10	Negative outcome wording letter	Positive = 1,805 Negative = 1,795	$F(2,9{,}854) = 22.12$, $p < 0.001$	Yes
Fieldwork compression	Tests two ways in which participants are asked to complete additional modules. sample members were allocated to one of five questionnaires; the continuous longer interview with a full set of rotating modules, the potential break-off request with a full set of rotating modules, the continuous longer interview with a reduced	13	Continuous longer interview with a full set of rotating modules Potential break-off request with a full set of rotating modules Continuous longer interview with a reduced set of rotating modules	Continuous longer interview, full set = 894 Break-off request, full set = 862 Continuous longer interview, reduced set =863 Break-off request, reduced set = 860 Control = 845	$F(5,9{,}854) = 6.15$, $p < 0.001$	Yes

	set of rotating modules, the potential break-off request with reduced set of rotating modules, or the control group.		Potential break-off request with reduced set of rotating modules			
Incentives	A number of incentive experiments were carried out, but fall under four categories; whether the respondent received a £10 unconditional incentive, a £10 unconditional plus £20 conditional, a £20 unconditional, or a £30 unconditional incentive.	All waves	No control or treatment group.	Differed throughout waves. Wave 2: £10 = 732 £5 = 1,412 Wave 16: £20 unconditional = 4,764 £30 unconditional = 883	$F(6,9{,}854) = 51.77$, $p < 0.001$	Yes
Mixed-modes	There have been a number of mixed mode experiments throughout the IP. For each interview, sample members were either in the face-to-face, web, or nurse mode.	All waves	No control or treatment group.	Differed throughout waves. Wave 1 =	$F(3,9{,}854) = 793.83$, $p < 0.001$	Yes

Appendix H

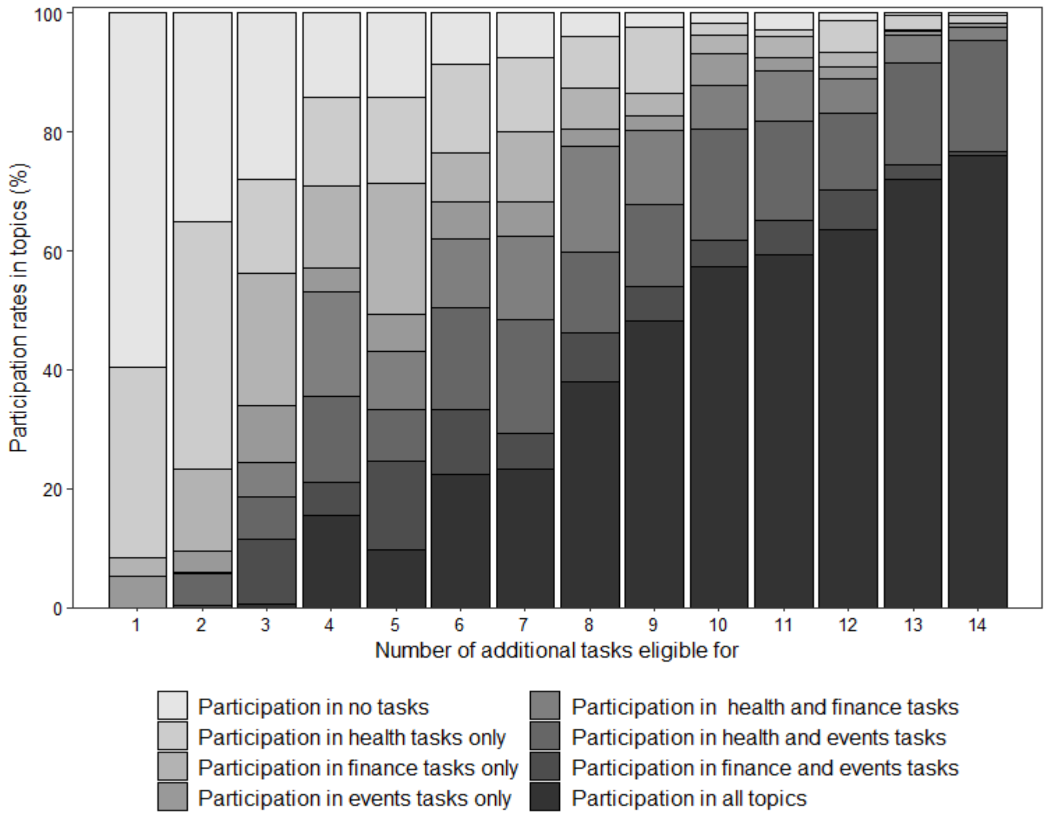
Table 8

Number of additional tasks Innovation Panel sample members participated in, by eligibility (IP6 – IP16).

Number of studies eligible for																												
1		2		3		4		5		6		7		8		9		10		11		12		13		14		N
N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
742	60	425	35	97	28	46	14	36	14	25	9	16	8	8	4	7	2	5	2	5	3	3	1	1	0	1	0	1,401
502	40	605	50	141	41	94	29	92	36	67	23	40	19	25	12	33	11	21	7	6	3	9	4	5	1	5	1	1,656
2	0	184	15	99	29	113	35	65	25	76	26	42	20	45	22	39	13	27	9	21	12	17	7	14	4	8	2	755
-	-	-	-	11	3	66	20	44	17	61	21	46	22	39	19	55	19	31	10	28	16	41	16	30	9	11	3	461
-	-	-	-	-	-	8	2	16	6	41	14	34	16	46	22	42	14	46	16	27	15	36	14	29	8	22	6	348
-	-	-	-	-	-	-	-	2	1	10	3	17	8	25	12	42	14	50	17	25	14	42	16	50	14	32	9	297
-	-	-	-	-	-	-	-	-	-	9	3	11	5	8	4	38	13	35	12	27	15	24	9	60	17	34	10	243
-	-	-	-	-	-	-	-	-	-	-	-	5	2	8	4	25	9	38	13	18	10	28	11	50	14	35	10	211
-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	1	7	2	29	10	7	4	25	10	42	12	55	16	165
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	2	11	4	4	2	15	6	26	7	41	12	104
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	1	5	3	5	2	21	6	27	8	61
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	1	9	4	14	4	34	10	60
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0	5	1	27	8	33
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	1	12	3	14
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	1	3
1,246	100	1,214	100	348	100	327	100	255	100	289	100	211	100	206	100	294	100	296	100	175	100	255	100	349	100	347	100	5,812

Appendix I

Figure 6. Participation rates in types of task topics in the Innovation Panel (IP6 –IP16), by eligibility.



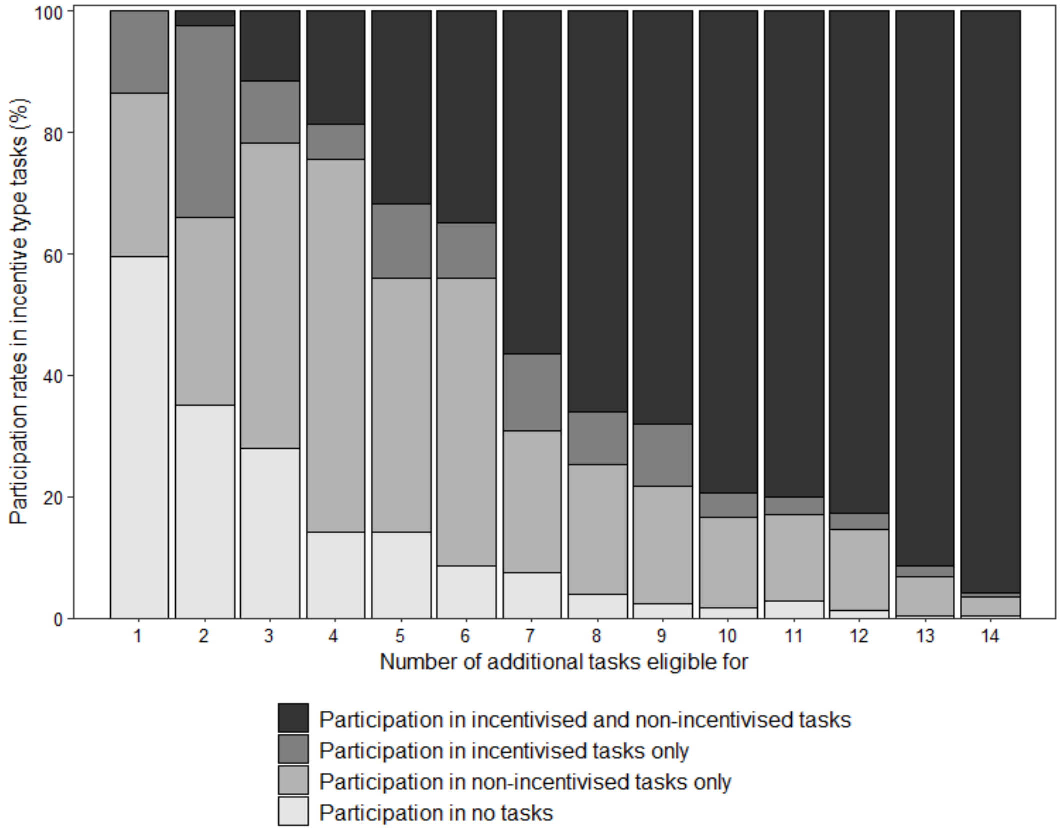
Note. Health tasks = Finger Measurement study, pre-interview blood pressure, hair sample, blood sample, Wellbeing app, BVI app, spatial navigation game.

Finance-related tasks = FCA consent, SS1, HMRC consent, SS2.

Event tasks = Time Diary, Life Events Study, SMS text consent.

Appendix J

Figure 7. Participation rates in incentivised and non-incentivised tasks in the Innovation Panel (IP6 –IP16), by eligibility.



Appendix K

Table 9

Participation pattern of additional tasks in the Innovation Panel for sample members eligible for all tasks (IP6 – IP16; Finger Measurement Study, Time Diary, FCA consent, SS1, HMRC consent, SS2, pre-interview blood pressure, hair sample, blood sample, Life Events Study, text consent, Wellbeing app, BVI app, Spatial Navigation Game).

Pattern	Freq.	%
111.1.111111..	6	1.73
111111.11111.1	5	1.45
111111111111.1	5	1.45
111.1..1111...	4	1.16
111.11111111.1	4	1.16
1.............	3	0.87
11............	3	0.87
11.......11...	3	0.87
11....1..11...	3	0.87
Other patterns	311	89.6
Total	347	100

Note. 1 = participated, . = did not participate

Appendix L

Table 10

Linear probability models of number of prior tasks invited to on participation in a task (IP6 – IP16).

	Unweighted		Weighted		Fixed effects		Item non-response control		Instrumental regression	
	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value
Number of tasks invited to previously	-0.001 (0.002)	0.729	-0.001 (0.003)	0.631	0.010 (0.004)	0.010	-0.001 (0.002)	0.540	-0.040 (0.003)	0.000
Task (HMRC consent)										
Finger measurement	0.277 (0.015)	0.000	0.268 (0.024)	0.000	0.304 (0.021)	0.000	0.524 (0.028)	0.000	0.174 (0.014)	0.000
Time diary	0.017 (0.015)	0.251	0.002 (0.022)	0.929	0.038 (0.020)	0.055	0.076 (0.017)	0.000	-0.069 (0.015)	0.000
FCA consent	0.019 (0.013)	0.149	0.012 (0.021)	0.570	0.030 (0.016)	0.052	0.018 (0.013)	0.179	-0.049 (0.013)	0.000
SS1	-0.423 (0.011)	0.000	-0.427 (0.019)	0.000	-0.461 (0.013)	0.000	-0.423 (0.012)	0.000	-0.453 (0.011)	0.000
SS2	-0.397 (0.011)	0.000	-0.404 (0.014)	0.000	-0.452 (0.013)	0.000	-0.393 (0.011)	0.000	-0.363 (0.011)	0.000
Blood pressure	-0.208	0.000	-0.251	0.000	-0.248	0.000	-0.175	0.000	-0.153	0.000

	(0.014)		(0.024)		(0.016)		(0.015)		(0.015)	
Hair sample	-0.186	0.000	-0.240	0.000	-0.246	0.000	-0.168	0.000	-0.092	0.000
	(0.014)		(0.023)		(0.018)		(0.015)		(0.015)	
Blood sample	-0.120	0.000	-0.149	0.000	-0.178	0.000	-0.091	0.000	0.013	0.456
	(0.015)		(0.024)		(0.020)		(0.016)		(0.017)	
Life Events	0.079	0.000	0.080	0.003	0.031	0.172	0.106	0.000	0.237	0.000
	(0.017)		(0.026)		(0.023)		(0.018)		(0.019)	
Text consent	0.086	0.000	0.066	0.010	0.029	0.214	0.138	0.000	0.259	0.000
	(0.015)		(0.025)		(0.023)		(0.017)		(0.019)	
Wellbeing app	-0.151	0.000	-0.163	0.000	-0.239	0.000	-0.103	0.000	0.066	0.002
	(0.017)		(0.023)		(0.027)		(0.019)		(0.021)	
BVI app	-0.443	0.000	-0.461	0.000	-0.581	0.000	-0.431	0.000	-0.204	0.000
	(0.018)		(0.023)		(0.030)		(0.019)		(0.021)	
Spatial Navigation Game	-0.158	0.000	-0.180	0.000	-0.239	0.000	-0.141	0.000	0.100	0.000
	(0.019)		(0.024)		(0.033)		(0.020)		(0.022)	
Sample Origin (original sample)										
IP4 refreshment sample	0.009	0.345	0.013	0.342	-	-	0.010	0.272	-	-
	(0.009)		(0.014)				(0.009)			
IP7 refreshment sample	0.00	0.977	0.009	0.557	-	-	-0.001	0.903	-	-
	(0.009)		(0.015)				(0.010)			
IP10 refreshment sample	0.114	0.000	0.130	0.000	-	-	0.103	0.000	-	-
	(0.017)		(0.024)				(0.018)			

IP11 refreshment sample	0.114	0.000	0.120	0.000	-	-	0.103	0.000	-	-
	(0.015)		(0.020)				(0.015)			
IP14 refreshment sample	0.240	0.000	0.230	0.000	-	-	0.227	0.000	-	-
	(0.022)		(0.029)				(0.023)			
Item non-response	-	-	-	-	-	-	-0.763	0.000	-	-
							(0.067)			
Mode (CAWI)										
CATI	-	-	-	-	-	-	0.061	0.000	-	-
							(0.008)			
CAPI	-	-	-	-	-	-	-0.112	0.000	-	-
							(0.015)			
Spending Study 1 incentive experiment (not in experiment)										
£2	0.004	0.714	0.020	0.266	-	-	-0.001	0.906	0.044	0.000
	(0.012)		(0.018)				(0.013)		(0.013)	
£6	0.147	0.000	0.155	0.000	-	-	0.138	0.000	0.184	0.000
	(0.013)		(0.016)				(0.013)		(0.013)	
HMRC consent question experiment (not in experiment)										
Standard consent question	0.014	0.120	0.011	0.390	-	-	-0.023	0.022	0.021	0.015
	(0.009)		(0.012)				(0.010)		(0.009)	
Easy consent question	0.026	0.004	0.023	0.073	-	-	-0.008	0.405	0.032	0.000

	(0.009)		(0.013)				(0.010)		(0.009)	
Spending Study 2 invitation experiment (not in experiment)										
Invited interwave	-0.001	0.905	-0.022	0.111	-	-	0.003	0.768	0.043	0.000
	(0.010)		(0.014)				(0.011)		(0.011)	
Invitation to app in-interview	-0.021	0.037	-0.042	0.004	-	-	-0.018	0.093	0.028	0.009
	(0.010)		(0.014)				(0.011)		(0.011)	
Blood pressure advanced letter experiment (not in experiment)										
Information Treatment	-0.076	0.008	0.095	0.026	-	-	0.111	0.002	-0.063	0.019
	(0.028)		(0.042)				(0.035)		(0.027)	
Pro-social appeal treatment	-0.080	0.005	0.101	0.015	-	-	0.108	0.002	-0.066	0.013
	(0.028)		(0.041)				(0.036)		(0.027)	
Control	-0.091	0.001	0.084	0.050	-	-	0.095	0.007	-0.078	0.004
	(0.028)		(0.042)				(0.035)		(0.027)	
Blood sample mode experiment (not in experiment)										
Face-to-face	-0.266	0.004	-0.034	0.795	-	-	-0.062	0.477	-0.174	0.014
	(0.094)		(0.131)				(0.087)		(0.071)	
Web	-0.200	0.003	-0.134	0.062	-	-	-0.056	0.461	-0.139	0.030
	(0.068)		(0.071)				(0.076)		(0.064)	
Nurse	0.038	0.119	0.047	0.183	-	-	0.065	0.007	0.058	0.015
	(0.024)		(0.035)				(0.024)		(0.024)	

Hair sample mode experiment (not in experiment)										
Face-to-face	0.322	0.000	-0.083	0.546	-	-	-0.079	0.395	0.267	0.000
	(0.087)		(0.137)				(0.093)		(0.066)	
Web	0.258	0.000	0.017	0.838	-	-	-0.049	0.553	0.235	0.000
	(0.066)		(0.082)				(0.083)		(0.062)	
Nurse	0.090	0.008	-0.089	0.103	-	-	-0.140	0.001	0.108	0.001
	(0.034)		(0.054)				(0.041)		(0.031)	
Text consent position experiment (not in experiment)										
Early	0.214	0.000	0.214	0.000	-	-	0.213	0.000	0.219	0.000
	(0.010)		(0.012)				(0.010)		(0.010)	
Late	0.228	0.000	0.222	0.000	-	-	0.229	0.000	0.235	0.000
	(0.009)		(0.012)				(0.009)		(0.009)	
Wellbeing app length experiment (not in experiment)										
10 minutes	-0.071	0.000	-0.059	0.000	-	-	-0.064	0.000	-0.065	0.000
	(0.011)		(0.015)				(0.011)		(0.011)	
2 minutes	-0.083	0.000	-0.068	0.000	-	-	-0.073	0.000	-0.080	0.000
	(0.010)		(0.016)				(0.011)		(0.010)	
BVI app incentive experiment (not in experiment)										
Unconditional £5 incentive	0.044	0.000	0.049	0.000	-	-	0.044	0.000	0.048	0.000

	(0.009)		(0.013)				(0.010)		(0.009)	
Conditional £5 incentive	0.060	0.000	0.066	0.000	-	-	0.061	0.000	0.065	0.000
	(0.009)		(0.011)				(0.009)		(0.009)	
Spatial Navigation Game incentive experiment (not in experiment)										
£10 conditional incentive	-0.040	0.000	-0.030	0.040	-	-	-0.042	0.000	-0.084	0.000
	(0.010)		(0.014)				(0.011)		(0.011)	
£30 conditional incentive	-0.041	0.000	-0.033	0.044	-	-	-0.042	0.000	-0.084	0.000
	(0.010)		(0.016)				(0.011)		(0.011)	
N	31,728	-	29,949	-	27,758	-	29,949	-	31,728	-

Appendix M

Table 11

Model fit statistics and regression coefficients with p-values across models with varying controls for the unweighted linear probability model (model 1) of participation in the next Innovation Panel additional task by the number of prior tasks invited to.

Measures of fit	Number of prior tasks	Task	Types of prior tasks invited to	Task experiments	Sample origin
Number of previous tasks β	-0.003	0.005	0.012	-0.015	-0.001
P-value	0.003	0.000	0.000	0.000	0.729
AIC	45,729.0	40,522.8	40,418.9	37,901.1	37,716.8
BIC	45,745.8	40,648.3	40,602.9	38,302.6	38,076.4
Included in final model?	-	Yes	No	Yes[1]	Yes

[1] Tasks experiments included if they provided a significant joint F-test.

Appendix N

Table 12

Linear probability models of number of prior tasks invited to on participation in the annual interviews (IP1 – IP16).

	Unweighted		Weighted		Fixed-Effects		Instrumental variable model	
	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value
Number of tasks invited to prior	0.086	0.000	0.069	0.000	0.015	0.000	-0.005	0.037
	(0.001)		(0.002)		(0.002)		(0.003)	
Interview (wave 15)								
1	1.480	0.000	0.871	0.000	1.263	0.000	0.783	0.000
	(0.019)		(0.146)		(0.031)		(0.019)	
2	1.555	0.000	0.717	0.000	1.292	0.000	0.934	0.000
	(0.021)		(0.053)		(0.030)		(0.027)	
3	1.296	0.000	0.710	0.000	1.029	0.000	0.672	0.000
	(0.021)		(0.051)		(0.030)		(0.027)	
4	1.359	0.000	0.697	0.000	1.030	0.000	0.753	0.000
	(0.020)		(0.049)		(0.029)		(0.026)	
5	0.755	0.000	0.605	0.000	0.405	0.000	0.120	0.000
	(0.016)		(0.029)		(0.025)		(0.021)	
6	0.908	0.000	0.647	0.000	0.542	0.000	0.265	0.000
	(0.015)		(0.030)		(0.024)		(0.021)	
7	0.721	0.000	0.518	0.000	0.392	0.000	0.145	0.000
	(0.014)		(0.028)		(0.022)		(0.019)	
8	0.637	0.000	0.530	0.000	0.345	0.000	0.127	0.000
	(0.015)		(0.025)		(0.022)		(0.018)	
9	0.733	0.000	0.534	0.000	0.442	0.000	0.235	0.000
	(0.018)		(0.029)		(0.025)		(0.021)	
10	0.376	0.000	0.405	0.000	0.019	0.628	0.005	0.754
	(0.025)		(0.024)		(0.039)		(0.016)	
11	0.480	0.000	0.364	0.000	0.267	0.000	0.096	0.000
	(0.013)		(0.022)		(0.018)		(0.016)	
12	0.192	0.000	0.210	0.000	0.036	0.045	-0.042	0.003

	(0.013)		(0.021)		(0.018)		(0.014)	
13	0.266 (0.066)	0.000	0.107 (0.022)	0.000	0.115 (0.090)	0.204	0.120 (0.023)	0.000
14	0.035 (0.007)	0.000	0.000 (0.011)	0.991	0.092 (0.010)	0.000	0.025 (0.008)	0.001
16	-0.001 (0.006)	0.832	0.004 (0.010)	0.654	0.046 (0.010)	0.000	0.030 (0.006)	0.000
Sample Origin (original sample)								
IP4 refreshment sample	0.120 (0.012)	0.000	0.065 (0.013)	0.000	-	-	-	-
IP7 refreshment sample	0.129 (0.013)	0.000	0.082 (0.016)	0.000	-	-	-	-
IP10 refreshment sample	0.290 (0.014)	0.000	0.217 (0.019)	0.000	-	-	-	-
IP11 refreshment sample	0.308 (0.013)	0.000	0.238 (0.020)	0.000	-	-	-	-
IP14 refreshment sample	0.734 (0.017)	0.000	0.591 (0.032)	0.000	-	-	-	-
Incentives (inapplicable)								
£10	-0.092 (0.015)	0.000	0.117 (0.142)	0.412	-0.096 (0.027)	0.000	-0.149 (0.018)	0.000
£10+£20	-0.027 (0.017)	0.106	0.147 (0.146)	0.315	-0.022 (0.029)	0.438	-0.059 (0.021)	0.005
£20	-0.099 (0.012)	0.000	0.132 (0.143)	0.358	-0.166 (0.025)	0.000	-0.175 (0.014)	0.000
£30	-0.006 (0.015)	0.692	0.157 (0.146)	0.284	0.009 (0.030)	0.761	-0.024 (0.021)	0.242
£5	-0.144	0.000	0.122	0.401	-0.142	0.000	-0.225	0.000

	(0.019)		(0.144)		(0.029)		(0.021)	
£15	-0.201	0.000	0.168	0.250	-0.204	0.000	-0.292	0.000
	(0.029)		(0.145)		(0.038)		(0.031)	
Mode (Inapplicable)								
Web	0.603	0.000	0.117	0.005	0.588	0.000	0.636	0.000
	(0.012)		(0.041)		(0.017)		(0.013)	
Face-to-face	0.607	0.000	0.111	0.007	0.569	0.000	0.635	0.000
	(0.012)		(0.040)		(0.017)		(0.013)	
Nurse	0.609	0.000	0.131	0.003	0.549	0.000	0.650	0.000
	(0.017)		(0.043)		(0.025)		(0.020)	
Individual mailings experiment (inapplicable)								
Sunday	0.375	0.000	0.126	0.177	0.133	0.317	0.535	0.000
	(0.098)		(0.093)		(0.133)		(0.101)	
Monday	0.557	0.000	0.104	0.155	0.245	0.027	0.736	0.000
	(0.063)		(0.073)		(0.111)		(0.064)	
Tuesday	0.352	0.000	-0.149	0.127	0.037	0.758	0.517	0.000
	(0.075)		(0.097)		(0.120)		(0.077)	
Wednesday	0.430	0.000	-0.041	0.650	0.193	0.074	0.589	0.000
	(0.064)		(0.089)		(0.108)		(0.065)	
Thursday	0.412	0.000	-0.079	0.413	0.072	0.520	0.593	0.000
	(0.067)		(0.097)		(0.111)		(0.070)	
Friday	0.432	0.000	-0.046	0.564	0.137	0.221	0.594	0.000
	-0.069		-0.079		-0.112		-0.071	
Saturday	0.515	0.000	0.235	0.021	0.253	0.048	0.666	0.000
	(0.085)		(0.100)		(0.128)		(0.085)	
Control	0.481	0.000	0.104	0.214	0.218	0.039	0.657	0.000
	(0.057)		(0.083)		(0.106)		(0.058)	
Unassigned	0.327	0.000	-0.086	0.201	0.180	0.070	0.406	0.000
	(0.050)		(0.067)		(0.100)		(0.050)	
Household mailings experiment (inapplicable)								
Sunday	-0.360	0.000	-0.065	0.542	-0.124	0.320	-0.532	0.000
	(0.087)		(0.106)		(0.125)		(0.087)	
Monday	-0.563	0.000	-0.102	0.207	-0.357	0.003	-0.757	0.000

	(0.071)		(0.080)		(0.120)		(0.072)	
Tuesday	-0.400 (0.073)	0.000	0.092 (0.083)	0.274	-0.144 (0.116)	0.216	-0.554 (0.075)	0.000
Wednesday	-0.366 (0.062)	0.000	0.103 (0.079)	0.195	-0.161 (0.104)	0.123	-0.520 (0.063)	0.000
Thursday	-0.325 (0.066)	0.000	0.110 (0.097)	0.259	-0.066 (0.107)	0.539	-0.493 (0.069)	0.000
Friday	-0.383 (0.076)	0.000	0.116 (0.075)	0.124	-0.113 (0.119)	0.341	-0.545 (0.077)	0.000
Saturday	-0.470 -0.091	0.000	-0.216 -0.113	0.059	-0.349 -0.134	0.009	-0.607 -0.092	0.000
Control	-0.422 (0.057)	0.000	-0.093 (0.085)	0.277	-0.227 (0.105)	0.032	-0.593 (0.058)	0.000
Unassigned	-0.261 (0.052)	0.000	0.129 (0.066)	0.052	-0.101 (0.101)	0.318	-0.334 (0.053)	0.000
Advance letter wording experiment (inapplicable)								
Positive wording	0.158 (0.026)	0.000	0.010 (0.017)	0.571	0.317 (0.039)	0.000	0.150 (0.022)	0.000
Negative wording	0.148 (0.026)	0.000	-	-	0.298 (0.039)	0.000	0.139 (0.022)	0.000
Compression experiment (inapplicable)								
Continuous interview, full set	-0.122 (0.066)	0.067	0.004 (0.026)	0.878	-0.011 (0.091)	0.906	-0.087 (0.025)	0.000
Break point request, full set	-0.194 (0.066)	0.003	-0.027 (0.028)	0.342	-0.109 (0.091)	0.229	-0.172 (0.025)	0.000
Continuous interview, reduced set	-0.105 (0.067)	0.113	0.010 (0.026)	0.696	-0.014 (0.091)	0.878	-0.072 (0.025)	0.004
Break point request, reduced set	-0.147 (0.067)	0.027	-0.040 (0.029)	0.175	-0.048 (0.091)	0.598	-0.122 (0.025)	0.000

Control	-0.153 (0.067)	0.021	-	-	-0.067 (0.091)	0.463	-0.102 (0.026)	0.000
Wave 2 mode experiment (inapplicable)								
Face-to-face	0.150 (0.012) (0.012)	0.000	0.096 (0.015)	0.000	-	-	0.14 (0.013)	0.000
CATI move one, move all	0.111 (0.012)	0.000	0.068 (0.016)	0.000	-	-	0.094 (0.014)	0.000
CATI try all	0.131 (0.012)	0.000	0.077 (0.018)	0.000	-	-	0.109 (0.013)	0.000
AIC	-	67,494	-	-	-	38,398	-	-
N	-	61,178	-	44,534	-	41,154	-	61,178

Appendix O

Table 13

Model fit statistics and regression coefficients with p-values across models with varying controls for the unweighted linear probability model (model 1) of participation in the next Innovation Panel annual interview by the number of prior tasks invited to.

Measures of fit	Number of prior tasks	Interview year	Interview experiments	Sample origin
Co-efficient	0.026	0.058	0.062	0.086
P-value	0.000	0.000	0.000	0.000
AIC	84,074.1	75,306.19	70,381.19	67,493.76
BIC	84,092.15	75,459.56	70,894.52	68,026.03
Included in final model?	-	Yes	Yes[2]	Yes

[2] Annual interview experiments included if they provided a significant joint F-test.