SURVEY FUTURES

SURVEY DATA COLLECTION METHODS COLLABORATION

Working Paper 21:

Methods for within-household selection in self-administered push-to-web surveys: an experimental comparison

Nathan Reece[1], Peter Lynn[2]

[1] City St. George's, University of London; [2] University of Essex

July 2026

www.surveyfutures.net

Acknowledgements

Survey Futures is an Economic and Social Research Council (ESRC)-funded initiative (grant ES/X014150/1) aimed at bringing about a step change in survey research to ensure that high quality social survey research can continue in the UK. The initiative brings together social survey researchers, methodologists, commissioners and other stakeholders from across academia, government, private and not-for-profit sectors. Activities include an extensive programme of research, a training and capacity-building (TCB) stream, and dissemination and promotion of good practice. The research programme aims to assess the quality implications of the most important design choices relevant to future UK surveys, with a focus on inclusivity and representativeness, while the TCB stream aims to provide understanding of capacity and skills needs in the survey sector (both interviewers and research professionals), to identify promising ways to improve both, and to take steps towards making those improvements. Survey Futures is directed by Professor Peter Lynn, University of Essex, and is a collaboration of twelve organisations, benefitting from additional support from the Office for National Statistics and the ESRC National Centre for Research Methods. Further information can be found at www.surveyfutures.net.

The research reported here forms part of the within-household selection subproject in Survey Futures Research Strand 4: Methods for surveys without field interviewer, led by Professor Olga Maslovskaya.

The survey conducted was designed at City St. George's, University of London in collaboration with the University of Essex, funded as a Survey Futures Phase 2 research project by UKRI and ESRC [grant number ES/X014150/1]. The survey platform used was managed and hosted by Centerdata in Tilburg, Netherlands. Data collection was undertaken by Opinion42.

Citation

Prior to citing this paper, please check whether a final version has been published in a journal. If so, please cite that version. In the meanwhile, the suggested form of citation for this working paper is:

Reece N. & Lynn P. (*2026*) 'Methods for within-household selection in self-administered push-to-web surveys: an experimental comparison', *Survey Futures Working Paper* no. 21. Colchester, UK: University of Essex. Available at https://surveyfutures.net/working-papers/.

Contents

Abstract....4
1. Introduction....4
2. Data and methods....6
3. Results....10
Response rates....10
Design effects....11
Relative cost....12
Compliance with random selection....12
Fraud in the any two sample....13
Sample composition....14
Attitudes....16
4. Conclusion and discussion....16
5. References....19

Methods for within-household selection in self-administered push-to-web surveys: an experimental comparison

Abstract

For self-administered postal push-to-web surveys in countries where only address frames are available, it is challenging to devise a method for within-household respondent selection that has both desirable statistical properties and high compliance levels. Several methods are currently used, with no consensus about their relative merits. The aim of this research is to provide much-needed evidence about the relative effectiveness of alternative methods of within-household selection in the absence of an interviewer.

Two common approaches, the any two method (non-random with potential cost advantages) and the next-birthday method (quasi-random), are compared via an experimental design. A factorial experiment was incorporated into the experimental design comparing two common styles of invitation letters. Research questions address differences between the methods in population representation, precision of estimation, and costs.

Overall, accounting for design effects, the any two method is cheaper to implement with similar survey outcomes. Little variation in precision of estimates or the distribution of answers to attitudinal questions is observed. A slightly higher proportion of younger respondents is observed in achieved sample of the next-birthday group. The average educational attainment is consistent across within household selection approaches, though this appears to be moderated by the type of invitation letter used.

Keywords: self-completion, sampling methods, experimental comparison, survey methods

1. Introduction

In many countries, the use of postal self-completion approaches for general population social surveys was rare prior to the covid-19 pandemic of 2020. With self-completion approaches, the absence of field interviewers poses challenges for delivering high quality surveys, as the burden of understanding and executing complex tasks falls instead upon the sample member. These tasks include selecting respondents within a household and persuading people to participate. Survey methodologists must make careful design decisions that take into account the abilities and motivations of potential respondents.

With sample frames of addresses it is hard to control the selection of a target respondent at an address, in the absence of an interviewer. Survey agencies and methodologists have developed and administered several approaches, with no agreement and very limited evidence on the relative merits of each. Many self-completion web surveys in the UK utilise a non-random within-household selection method despite describing themselves as probability-based, such as requesting any two adults in a household to participate (Ndebele et al., 2026 forthcoming). It is unclear whether this introduces bias towards the most willing persons (Koch, 2018). This decision is mostly driven by costs rather than quality but such methods are also practical as it is difficult to design a method that facilitates random selection without an interviewer present. Random methods, such as the next-birthday method, are used in some surveys. There is little evidence of whether implementing a non-random sampling method at this final stage has significant impacts on the results and representation of the survey, in terms of survey estimates, sample composition, or response rates.

A study on the Community Life Survey (Williams, 2019) compared results when asking all adults to complete the survey with a simulated random selection of one person, performed after the survey concluded. This demonstrated few significant differences in attitudinal variables. However, the report acknowledged that its conclusions were not robust as they were unable to fairly simulate the implementation of non-random and random methods. The analysis was also unable to compare sample composition or response rates between the random-one condition versus the non-random conditions, as the former was simulated rather than empirically tested.

There is no direct evidence on which within-household selection method provides better representation or higher response rate. In particular, requesting more than one person per household to respond could on the on hand be seen as more burdensome to a household, but on the other hand avoids the need for complex instructions for conducting random selection of one person. No survey experiment to address this issue has been conducted in the UK or elsewhere.

The most common random methods for within-household selection in self-completion surveys are birthday methods (Nicolaas, 2022). These can be implemented online and on paper surveys (European Social Survey, 2024). It is known that random birthday methods in self-completion surveys do not always result in the correct person participating (Olson et al., 2014; Olson & Smyth, 2017). Little is known, however, about whether non-compliance affects representation (as it is difficult to ascertain when the correct person has responded).

Respondents in postal self-administered web and paper surveys may be informed of who within their household should complete the questionnaire in the invitation letter, subsequent reminders, and the questionnaire itself. Check questions in the questionnaire have been shown to improve compliance levels with the next-

birthday method (Olson & Smyth, 2017). The relative success of a within-household selection method is therefore likely to be sensitive to whether the instructions are read carefully, how well they are understood, and how the importance of following the instructions correctly is perceived – all of which makes the careful design of invitation letters an important focus for survey methodologists.

To assess whether the implementation of different within-household selection methods may affect the outcomes of a self-completion web and paper survey, an experiment was conducted among the adult population in Great Britain. Two common approaches were implemented, the any two method (non-random with potential cost advantages) and the next-birthday method (quasi-random). Survey outcomes of response rates, sample composition, compliance with random selection, standard errors, and responses to attitudinal variables were examined, in addition to the relative cost of each approach. To this end, the following research questions are answered:

1. Does either within-household selection method's response rate significantly outperform the other?
2. Are the survey estimates of one method more precise than the other?
3. Does the cost of each approach differ?
4. Does either approach result in a more representative sample?
5. To what extent is each method compliant with random selection?
6. Are there significant differences in the distribution of responses to attitudinal questions?

The survey also incorporated a factorial experiment where the invitation and reminder letters varied. The extent to which answers to the above research questions was moderated by the letter version used was also evaluated.

2. Data and methods

A self-completion survey in Great Britain (England, Scotland, and Wales) was conducted among a target population of all adults aged 18 and over, using a reduced version of the European Social Survey Round 12 questionnaire (European Social Survey, 2026). The questionnaire aims to measure political and social attitudes relating to a wide array of topics such as trust, wellbeing, health, immigration. A stratified random sample of 5156 addresses was selected from the Postal Address File (PAF), with the probability of selecting each address being equal. The postal survey was fielded by a survey agency in collaboration with the University of Essex and City St George's University between 15 May and 31 July 2025. The overall household-level response rate was 35.0 percent (AAPOR RR1, AAPOR 2023), with at least one complete or partial questionnaire received from

1,805 households. In total, 2,152 complete or partial questionnaires were received, of which 1,681 were completed on web and 471 on paper.

The next-birthday method was employed for a random subset of 3,008 addresses, and the any two method was used for the remaining 2,148 addresses. These sample sizes were chosen in order that each method should produce similar numbers of respondents and thereby maximise statistical power for comparisons of the two responding samples. The next-birthday method is a quasi-probability selection method that is commonly used in postal surveys (Koch, 2018; Smyth, Olson, & Stange, 2019). The any two method is a non-probability method used commonly in the UK in postal self-completion surveys (Ndebele et al., 2026 forthcoming). For households with two or fewer eligible persons the probability of selection within the household is 100%, however for households with greater than two eligible persons selection is not random and therefore the probability of selection is unknown.

An invitation letter providing access to the web questionnaire was sent to all households, followed by three reminder letters that were sent to all nonresponding households aside from any that had opted out of the survey by reaching out to the helpline. Households in each within-household selection treatment group received a singular £5 unconditional incentive with the invitation letter (bank note), and a £10 voucher was promised to each respondent who completed a questionnaire. The second reminder letter included one paper questionnaire in the next-birthday group. For the any two group, if no web questionnaire had been completed by the time the 2nd reminder mailing was being prepared, two paper questionnaires were sent. If one web questionnaire was completed, one paper questionnaire was sent alongside a letter acknowledging the receipt of one completed questionnaire and requesting another. The 3rd and final reminder sent to the any two group was similarly modified for households where one completed questionnaire had already been received (but without the inclusion of any paper questionnaire).

The two within-household selection methods were fully crossed with two randomly assigned versions of invitation and reminder letters. One version of the invitation letters was composed in line with social exchange theory as described in *the Tailored Design Method* (Dillman, Smyth, & Christian, 2014), whereas the design of the other version was grounded in the respondent centred design framework as described in *Respondent Centred Surveys* (Wilson & Dickinson, 2021). Hereafter, the former letter version is referred to as the 'TDM letter' and the latter as the 'RCS letter'. The four types of invitation letters are shown below in figure 1 and figure 2.

Figure 1: Example invitation letters for the next birthday method. From left to right, invitation letter style based on respondent centred design framework (RCS letter), invitation letter style based on social exchange theory/the tailored design method (TDM letter).

Figure 2: Example invitation letters for the any two method. From left to right, invitation letter style based on respondent centred design framework (RCS letter), invitation letter style based on social exchange theory/the tailored design method (TDM letter).

The questionnaire content across the two within-household selection methods was identical aside from the presence of instructions for the within-household selection method on the cover page of the paper questionnaire and the presence of a next-birthday check question at the beginning of the web and paper questionnaire in the next-birthday group as recommended by Olsen & Smyth (2017) to boost compliance with the method.

All analysis for this paper is based on the set of partial and complete questionnaires (n=2,152). Complete questionnaires were defined as all questionnaires where 80 percent or more of questions asked to all respondents were answered. Partial questionnaires were defined as all questionnaires where 50 percent or more but less than 80 percent of questions asked to all respondents were answered.

Response rate analyses are unweighted as the probability of selection of each address was equal. Chi-square tests are used to evaluate differences at the household level. Individual level response rates are then compared relying on an external estimate of the average number of eligible persons per household in the any two group.

Variable costs for each within-household selection method group were calculated by adding the direct costs of postage, printing, and incentives. It should be noted that the cost per respondent depends on response rate and speed of response (as these determine the number of reminders that need to be sent, and response rate also determines the ratio of unconditional to conditional incentive payments) as well as the mean number of persons per household eligible to complete the questionnaire (which is higher with the any-two method).

To evaluate which method is more cost efficient in terms of the ratio of cost to effective sample size, we require estimates of design effects for each method. Design effects due to differences in selection probabilities were estimated based on the observed sample distributions of probabilities (which are expected to show less variation with the any-two method). For the next-birthday group, within household selection probabilities were assumed to be $1/N_i$, where N_i indicates the number of persons aged 18 or over in household i, whereas for the any two method, the within-household selection probability was assumed to be $2/N_i$ if $N_i \geq 3$ and 1 otherwise. Design effects due to clustering of attitudes within households were by definition non-existent (equal to 1.0) for the next-birthday method but for the any two group were estimated as follows. Intra-cluster correlation coefficients (ICCs) were calculated for each question asked to all respondents and the average of all individual ICCs was used for between-method comparisons. Alternative calculations of ICCs using different lists of variables (e.g., excluding household-level or socio-demographic variables) were also calculated.

Response rates are compared overall and within regions and Index of Multiple Deprivation (IMD) tertiles. Sample composition of each within-household selection group is compared using demographic characteristics of sex, age,

education, and citizenship. These characteristics are compared individually to benchmark census estimates collated from the UK Office of National Statistics and the National Records of Scotland. Only three categories of educational attainment were analysed as, taking into account how educational attainment was measured within the survey, this is the only categorisation for which comparable population estimates are available. The results of Wald chi-square tests of association are presented to assess whether statistically significant associations between the experimental conditions and demographic variables are observed.

Differences in a select group of thirteen attitudinal questions that span different topics were assessed, including climate change, internet usage, social trust, trust in institutions, satisfaction with the government, left/right ideology, gay and lesbian rights, wellbeing, immigration, and religiosity. The precision of estimates was also evaluated across the two within-household selection groups.

The extent to which the letter experiment moderated any of the observed impacts (or lack thereof) of the within-household selection method was assessed for each of the above topics.

3. Results

Response rates

Household-level response rates for each of the experimental conditions are shown below in table 1. Response rates across the two within-household selection methods were not statistically significant overall (chi-sq F statistic = 0.03, p = 0.861), when restricted to the TDM letter group (chi-sq F statistic = 0.06, p = 0.803), nor when restricted to the RCS letter group (chi-sq F statistic = 0.00, p = 0.996). The overall household-level response rate of the TDM letter group (AAPOR RR2 = 37.5%) was significantly higher than the RCS letter group (AAPOR RR2 = 32.6%) at the 99% confidence level (chi-sq F statistic = 13.34, p = 0.000).

Table 1: Household-level response rates by within-household selection method and letter conditions

Within-household selection method	Letter version	Total number of addresses	Households with at least one complete or partial	AAPOR RR2 (household level)
Next-birthday	TDM	1504	566	37.7%
Any two	TDM	1074	399	37.2%
Next-birthday	RCS	1504	490	32.6%
Any two	RCS	1074	350	32.7%

The overall AAPOR RR2 response rate for the next-birthday group was 35.1%, which is equivalent to the individual-level response rate since only one person per

household was selected. Using an external estimate of 1.655 for the average number of eligible adults aged 18 or over in each household in Great Britain with the any two method (Ndebele et al., 2026, forthcoming), the estimated AAPOR RR2 individual-level response rate for the any-two group was 30.9%, significantly lower than that of the next-birthday group at the 99% confidence level (chi-sq F statistic = 13.36, p = 0.000). The proportion of multi-adult responding households in which two questionnaires were completed was 68.5% with the TDM letter and 59.2% with the RCS letter.

Overall, 1,681 questionnaires were at least partially completed online and 471 questionnaires were at least partially completed on paper. Table 2 below shows the mode breakdown for each of the four experimental conditions. The TDM letters resulted in a significantly higher proportion of paper questionnaires being completed (chi-sq F-statistic = 8.73, p = 0.003).

Table 2: Breakdown of completions by mode by within-household selection method and letter conditions (row %)

Within-household selection method	Letter version	Web partials and completes	Paper partials and completes	Complete and partial interviews
Next-birthday	TDM	76.1%	23.9%	566
Any two	TDM	75.2%	24.8%	600
Next-birthday	RCS	81.8%	18.2%	490
Any two	RCS	80.2%	19.8%	496

Design effects

Table 3 below displays the calculated overall design effects (DEFF) for each within-household selection method. Overall, design effects are larger in the any two group due to within household clustering of attitudes. The design effect due to within household clustering compensates for a comparatively lower design effect due to variation in selection probabilities, resulting in a larger overall estimate. The larger design effect in the any two group indicates that a larger net sample of interviews (about 10% larger) is needed to achieve the same net effective sample size (NEFF) as the next-birthday group.

Table 3: Overall design effects and net effective sample sizes by within-household selection method

Within-household selection method	DEFF due to selection probabilities	DEFF due to within-household clustering	DEFF	Completes & partials	NEFF
Next-birthday	1.176	NA	1.176	1,056	898
Any two	1.084	1.195	1.295	1,096	846

The any two DEFF presented in table 3 (1.295) was calculated by determining the intra-cluster correlation coefficient for each question asked to all respondents. Alternate calculations were also made with factual household level questions removed (DEFF = 1.287), factual and subjective household level questions removed (DEFF = 1.285), factual household level questions and socio-demographic questions removed (DEFF = 1.279) and with factual/subjective household level questions and socio-demographic questions removed (DEFF = 1.277).

Relative cost

Across the two within-household selection methods, variable costs included unconditional incentives, packaging, printing, and postage. In the any two group incentive costs were saved as only one £5 note was included with the singular invitation letter sent to each sampled address, even though two people could ultimately fill in the survey. Conditional incentive costs per completed or partial questionnaire were roughly equivalent across the approaches as a conditional voucher worth £10 was sent to each respondent that provided contact details to receive it. Conditional incentives were sent to 873 respondents in the next-birthday group and to 971 respondents in the any two group. In both groups, 76 of the conditional incentives were sent physically with the rest sent digitally via email; each physical incentive costed £11 to disburse whereas each digital incentive costed £10. Table 4 below shows the breakdown of each of these costs and the computed cost per effective usable complete or partial interview. Overall, the average cost per effective interview is around £4.50 more in the next-birthday group (an 11% greater cost), with the cost of postage and incentives contributing a similar amount of additional cost. Using a more restrictive DEFF calculation excluding household level or socio-demographic variables would yield more savings for the any two method compared to the next-birthday method.

Table 4: Variable costs per effective usable questionnaire by within-household selection method

Within-household selection method	Total number of addresses	NEFF	Unconditional incentives	Conditional incentives	Print, Pack & Postage	Cost per effective interview
Next-birthday	3,008	898	£15,040	£8,816	£15,791	£44.14
Any two	2,148	846	£10,740	£9,776	£13,044	£39.68

Compliance with random selection

The any-two method is a random probability selection method only for households with fewer than three eligible residents. In the UK, 72% of households have only one or two persons aged 18 or over (Ndebele et al., 2026 forthcoming); in the other 28% of households some of the survey respondents would not have been selected had a random selection method been used.

Whether wittingly or not, the next-birthday method is not always implemented correctly by members of sample households. In our study, the questionnaire requested full dates of birth of the respondent and other household members. By reconciling the date the survey was filled in[1] with the reported birthdates, it was possible to assess how frequently respondents in the next-birthday group were indeed the household member 18 or older who had the next-birthday at the time that the questionnaire was completed. Overall, we estimate that 67.1% of respondents had the next birthday in their household, 14.8% did not, and in the remaining 18.1%% of cases it was unclear whether or not the respondent had the next birthday due to missing data or tied birthdays[2]. Amongst households with two or more eligible people, 54.8% of respondents had the next birthday, 21.2% did not, and 24% were inconclusive. Compliance with the next-birthday method was higher where the TDM letter was used (69.3% correct overall, 59.6% in multi-adult households) compared to where the RCS letter was used (64.7% correct overall, 49.1% correct in multi-adult households).

Fraud in the any two sample

The presence of fraud (i.e., where the same individual has responded twice in the same household even though they are aware they should only answer once) in the any two sample was assessed by analysing paradata in the form of hashed IP addresses and user agents, the presence of contact details for administering incentives, duration and timing of web survey completion, and responses to the questionnaire in households where more than one survey response was received. Overall, the number of cases where fraud appeared to be plausible was negligible (fewer than 5 households).

A higher proportion of respondents in the any two sample (88.4%) provided contact information to receive the conditional incentive than in the next-birthday sample (82.6%). This difference was statistically significant at the 99% confidence level (chi-sq F statistic = 10.91, p = 0.001). This may be due to the presence of otherwise unobserved fraud or because individuals within a household who might be motivated by the incentive to participate have the opportunity to do so with the any two method, but not with next birthday.

[1] For web respondents, the day in which they started the survey was used as the date of selection as the survey begins with a check question asking if the respondent has the next birthday. For paper respondents provided date of completion was taken. For 19 paper respondents who did not provide a date, the modal date of 17 June 2025 was imputed, which is also the date when the vast majority of paper questionnaires would have arrived in the post.

[2] 14.2% of all households could not be classified because respondents did not provide enough birthdate data for other household members or did not answer questions about overall household size. 0.6% of households reported that the person with the next birthday shared their birthday with another eligible household member. A further 3.3% of households reported that another eligible household member had the exact same birthdate as they did, which may suggest that they erroneously filled in their own birthdate twice.

Sample composition

Assessing response rate variation by population subgroups is restricted to region and IMD tertiles, as these are the only indicators available on the sampling frame. No significant differences between the two within household selection methods are observed in household-level response rates within region or IMD tertile. This finding was not moderated by the letter experiment. Table 5 shows the sex, age, educational attainment, and citizenship breakdowns of the achieved sample for each experimental condition.

Table 5: Sample composition of achieved samples in each experimental condition

	Percentages for experimental conditions (95% CI in parentheses)					**Wald chi-sq tests of association**					
	Next-birthday, TDM letter (*n* = 566)	**Next-birthday, RCS letter (*n* = 490)**	**Any two, TDM letter (*n* = 600)**	**Any two, RCS letter (*n* = 496)**	**Overall (*n* = 2152)**	***All 4 conditions***		***Within-household selection only***		***Letter version only***		**UK census data**
	% (95% CI)	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*	*F*	*p-value*	*F*	*p-value*	*F*	*p-value*	
Sex												
Female	55.7 (51.2, 60.2)	51.3 (46.5, 56.2)	54.1 (51.1, 57.2)	54.8 (51.2, 58.4)	53.9 (51.7, 56.2)	0.6	0.61	0.2	0.7	1.2	0.28	51.6
Education												
No qualifications (ISCED <1)	10.43 (8.1, 13.3)	10.0 (7.6, 13.1)	10.1 (7.9, 12.9)	10.7 (8.0, 14.3)	10.3 (9.0, 11.8)	1.5	0.16	0.6	0.54	1.3	0.27	17.8
Lower or upper secondary qualifications (ISCED = 2/3)	37.9 (33.6, 42.4)	46.5 (41.6, 51.4)	41.1 (36.6, 45.8)	37.2 (32.8, 41.9)	40.9 (38.6, 43.4)							46.5
Post-secondary qualification (ISCED 4+)	51.7 (47.1, 56.2)	43.5 (38.7, 48.4)	48.7 (44, 53.5)	52.0 (47.1, 56.9)	48.8 (46.3, 51.2)							35.7
Age												
18–24	7.2 (4.8, 10.6)	7.5 (4.8, 11.4)	5.9 (3.9, 9)	2.9 (1.6, 5.3)	6.3 (5, 7.9)	2.1	0.005	1.1	0.34	3.5	0.002	10.5
25–34	14.9 (11.9, 18.4)	7.6 (5.4, 10.5)	14.6 (11.2, 18.9)	10.5 (7.7, 14.1)	11.9 (10.4, 13.6)							17.0
35–44	14.6 (11.7, 18)	13.4 (10.5, 17)	13.7 (10.6, 17.4)	13.0 (10, 16.7)	13.8 (12.2, 15.5)							16.9
45–54	16.4 (13.2, 20.2)	20.2 (16.4, 24.6)	18.0 (14.5, 22.1)	16.4 (12.7, 20.8)	17.9 (16.0, 19.9)							15.5
55–64	15.8 (12.8, 19.5)	20 (16.3, 24.3)	16.3 (13.1, 20.1)	25.1 (20.8, 30)	18.8 (16.9, 20.8)							16.2
65-74	17.4 (14.3, 21)	16.3 (13.2, 20)	18.0 (14.5, 22.1)	19.0 (15.4, 23.3)	17.5 (15.8, 19.4)							12.1
75+	13.7 (11.0, 16.9)	15 (12.2, 18.4)	13.5 (10.6, 17.1)	13.2 (10.2, 16.9)	14.0 (12.5, 15.6)							11.8
Citizenship (among ages 18-64)												
% Citizens	87.5 (83.7, 91.3)	87.1 (82.8, 91.4)	88.7 (84.6, 92.8)	86.7 (82.6, 90.8)	87.5 (85.4, 89.6)	0.2	0.91	0.1	0.81	0.2	0.63	88.60
Average absolute deviation from census	4.4	4.3	4.3	6.5								

Sex, and citizenship were not significantly associated with any of the four experimental conditions. The proportion of women in the next-birthday/RCS letter version condition was around 3-4 percentage points closer to the population benchmark compared to the three other conditions, but the difference compared to the other groups was not statistically significant.

A statistically significant association between age (in the form of the seven groups presented in table 5) and the four experimental conditions was observed. The chi-squared test conducted only for the within-household selection conditions was not statistically significant, whereas it was significant when conducted only for the letter versions, suggesting that the overall significant association is moderated by the letter version. From the estimates it is apparent that larger differences in the age profiles are apparent across the within-household selection approaches when only considering the RCS letter version. For example, the youngest age group of ages 18-24, a population subgroup often underrepresented in social surveys (Wolf et al., 2021), is better represented in the next-birthday group overall, as the population benchmark is within the 95% confidence interval for both the next-birthday/TDM letter version and the next-birthday/RCS letter version, whereas the population benchmark is outside the confidence interval for both complementary estimates from the any two group. However, the difference between the next-birthday and the any two estimates is only statistically significant when restricting the analysis to the RCS letter version.

Overall, no statistically significant association was found between educational attainment and the four experimental conditions. The same finding was observed when expanding the educational attainment variable to seven categories (spanning ISCED 1-7+) which is not presented in table 5 above. Respondents with no qualifications (i.e., people who only have a primary level of educational attainment or less) were significantly underrepresented compared to the population as the benchmark statistic was well outside the 95% confidence intervals, and estimates across the four experimental conditions were very close to each other. Respondents with post-secondary qualifications were significantly overrepresented compared to the population for all four experimental conditions.

Respondents with lower or upper secondary qualifications were significantly underrepresented compared to the population in all experimental conditions aside from the next-birthday/RCS letter group. Further analysis of the achieved sample of the RCS letter version reveals that the next-birthday group had significantly more respondents with lower or upper secondary qualifications than the any two group when the RCS letter was used.

The absolute deviation between the survey estimates and the population benchmark was computed for each indicator presented in table 5, and then the average absolute deviation was computed. The any two/RCS letter version

performed the worst with the highest average absolute deviation, which at 6.5 was around 2 percentage points higher than the other three groups.

Attitudes

The mean response to thirteen attitudinal questions in the questionnaire across the experimental conditions, along with the corresponding standard errors, are displayed below in table 6.

Table 6: Mean responses to thirteen attitudinal questions in each experimental condition. Standard errors in parentheses.

	Next-birthday, TDM letter (*n* = 566)	**Next-birthday, RCS letter (*n* = 490)**	**Any two, TDM letter (*n* = 600)**	**Any two, RCS letter (*n* = 496)**	**Next-birthday (*n* = 1056)**	**Any two (*n* = 1096)**	**TDM letter (*n* = 1166)**	**RCS letter (*n* = 986)**	**Overall (*n* = 2152)**
Responsibility for climate change	6.3 (0.12)	6.4 (0.14)	6.1 (0.13)	6.4 (0.13)	6.3 (0.09)	6.2 (0.09)	6.2 (0.09)	6.4 (0.1)	6.3 (0.07)
Internet usage	4.6 (0.04)	4.5 (0.05)	4.5 (0.05)	4.5 (0.05)	4.5 (0.03)	4.5 (0.03)	4.5 (0.03)	4.5 (0.03)	4.5 (0.02)
General trust	4.7 (0.12)	4.9 (0.13)	4.9 (0.12)	5 (0.14)	4.8 (0.09)	4.9 (0.09)	4.8 (0.09)	4.9 (0.09)	4.8 (0.06)
Trust in legal system	4.7 (0.12)	4.5 (0.13)	4.8 (0.12)	4.9 (0.14)	4.6 (0.09)	4.8 (0.09)	4.7 (0.09)	4.7 (0.1)	4.7 (0.06)
Trust in police	5.6 (0.11)	5.6 (0.13)	5.7 (0.13)	5.6 (0.15)	5.6 (0.09)	5.6 (0.09)	5.6 (0.08)	5.6 (0.1)	5.6 (0.06)
Trust in political parties	2.4 (0.1)	2.5 (0.11)	2.3 (0.11)	2.7 (0.12)	2.4 (0.07)	2.5 (0.08)	2.4 (0.07)	2.6 (0.08)	2.4 (0.05)
Satisfaction with government	2.8 (0.1)	2.8 (0.11)	2.8 (0.12)	3.1 (0.13)	2.8 (0.08)	2.9 (0.09)	2.8 (0.08)	2.9 (0.09)	2.8 (0.06)
Left/Right scale	4.9 (0.1)	4.9 (0.11)	4.8 (0.11)	4.7 (0.11)	4.9 (0.07)	4.7 (0.08)	4.8 (0.08)	4.8 (0.08)	4.8 (0.05)
Life satisfaction	6.3 (0.1)	6.2 (0.12)	6.2 (0.11)	6.3 (0.12)	6.3 (0.08)	6.2 (0.08)	6.3 (0.08)	6.2 (0.09)	6.3 (0.06)
Gay/Lesbian right to adopt	2.1 (0.05)	2.3 (0.06)	2.2 (0.06)	2.3 (0.07)	2.2 (0.04)	2.2 (0.05)	2.1 (0.04)	2.3 (0.05)	2.2 (0.03)
Immigration bad/good for economy	4.9 (0.13)	4.8 (0.15)	4.9 (0.14)	5.3 (0.14)	4.8 (0.1)	5.1 (0.1)	4.9 (0.1)	5.0 (0.11)	4.9 (0.07)
How happy are you	6.5 (0.1)	6.5 (0.11)	6.4 (0.11)	6.6 (0.11)	6.5 (0.07)	6.5 (0.08)	6.5 (0.08)	6.6 (0.08)	6.5 (0.05)
Religiosity	3.3 (0.13)	3.5 (0.15)	3.1 (0.16)	3.6 (0.17)	3.4 (0.1)	3.3 (0.12)	3.3 (0.1)	3.5 (0.12)	3.4 (0.08)

Overall, only marginal differences were observed at the aggregate level between next-birthday and any two. No material difference was observed in standard errors between next-birthday and any two.

4. Conclusion and discussion

Survey outcomes from the two within-household selection methods were very similar in terms of sample composition, standard errors, and responses to attitudinal variables, though there was some moderation by the version of invitation/reminder letters administered. Response rates at the household level

were not significantly different. The individual response rate for the next-birthday method was significantly higher than for the any two method. The cost per effective completed questionnaire was around 12% higher with the next-birthday method, despite a lower design effect (the cost per completed questionnaire was 23% higher with next-birthday, ignoring design effects).

The Tailored Design Method letter significantly outperformed the Respondent Centred Surveys letter, resulting in higher response rates and significantly lower administration costs. Sample composition was also better overall compared to census data where the Tailored Design Method letter was used. To maximise response rates and minimise costs, then, the Tailored Design Method letter style would appear to be preferable. Where that style of letter is used, overall survey quality appears to be very similar with the two selection methods, so either would be acceptable. However, surveys costs to deliver estimates of equivalent precision are considerably higher with the next-birthday method, so the any two selection method can be safely implemented to reduce costs without any compromise on survey quality.

Our study has some limitations. It was carried out in the UK and there is no guarantee that findings would be similar in other countries. In particular, in countries with higher proportions of larger households, it can be expected that a higher proportion of respondents with the next birthday method will not in fact be the household member with the next birthday, and similarly, with the any two method a higher proportion of respondents will not be those who would have been selected with a random selection method. A greater departure from random selection of this kind could lead to less representative samples for either selection method.

The estimation of the design effect due to clustering in the any two sample is sensitive to the questionnaire fielded. This study used a pared down version of the European Social Survey Round 12 questionnaire, which measures attitudes and behaviours related to a wide array of topics. Certain question topics may result in more within-household clustering of responses than others, especially if they relate to experiences that may be shared among household members (e.g., how often one attends religious services). Furthermore, the questionnaire also included some household level questions (e.g., whether a household member has been victim of a burglary or assault) which are naturally clustered to a high degree at the household level. Surveys that aim to measure more household level indicators will therefore observe higher design effects which would reduce the cost-effectiveness of the any two method.

The findings of this study may also be sensitive to certain design decisions implemented to increase the overall quality of the survey and the administration of each method. For example, surveys that do not employ check questions in the survey to increase compliance with the next-birthday method or targeted

reminders to encourage a second response from multi-adult households may achieve poorer survey outcomes. Furthermore, the inclusion of a £5 cash unconditional incentive, which is uncommon in UK self-completion surveys, is considered by the authors to be an important factor in the relatively high response rates achieved. Had this method not been employed, it is likely that lower response rates would have been achieved overall which could have led to different patterns of nonresponse bias specific to each within household selection method.

Findings concerning the experimental test of invitation letter design are also limited to the particular design of the two letter versions. Both letters were designed by applying high-level principles from the respective approach to existing European Social Survey letters; they were not designed from first principles, involving user testing. The two letter versions varied in multiple dimensions such as text, layout, colour, and font face. The findings described herein cannot disentangle the individual effects of each varied element. Further research on the effects of these specific elements is needed in this area to better understand how they may impact response rates.

Further research on within household selection methods may look at the difference between the any two method and any three or any four methods. As more people are selected in each household, postage, and unconditional incentive costs are expected to go down. Printing costs could go either up or down, depending on the mailing protocol adopted. However, any cost savings may be offset by greater within household clustering which will increase design effects and therefore require larger sample sizes. Practical limitations may also become evident for providing equal access to paper questionnaires as is possible in the any two method. The risk of fraudulent responses completed by the same household member is also expected to rise with any three or any four methods. The relative performance of these methods could depend on the distribution of household size, so it would be informative to replicate experimental studies in countries with larger average household size than the UK.

Where the any two method is used, further research on the composition of the achieved sample with more than two adults residing in a household would also be beneficial to better understand response patterns for the part of the sample where selection is not random. In our study the sample size of this group was not large enough to make any robust observations.

5. References

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Wiley: Indianapolis, Indiana.

European Social Survey European Research Infrastructure (ESS ERIC) (2024). *ESS Round 12 Survey Specification for ESS ERIC Member, Observer and Guest Countries* (v2). London: ESS ERIC Headquarters

European Social Survey European Research Infrastructure (ESS ERIC) (2026) ESS round 12 - 2026. Personal and social wellbeing, Attitudes to immigrants and refugees. Sikt - Norwegian Agency for Shared Services in Education and Research.

Koch, A. (2018) Within-household selection of respondents. Chapter 5, p.93-111, in *Advances in Comparative Survey Methods* (ed.s T.P. Johnson, B.-E. Pennell, I.A.L. Stoop & B. Dorer), Wiley.

Nicolaas, G. (2022). Within-Household Selection for Push-to-Web Surveys. *GenPopWeb2 Report. Available at https://www. ncrm. ac. uk/documents/Within% 20household% 20selection% 20for% 20push% 20to% 20web% 20surveys. pdf*.

Ndebele N, Lynn P, Fitzgerald R & Comanaru R (2026, forthcoming) 'Within-household Selection Methods for Surveys without Field Interviewers in the United Kingdom: Evidence review', *Survey Futures Report*. Colchester, UK: University of Essex. Available at https://surveyfutures.net/reports/.

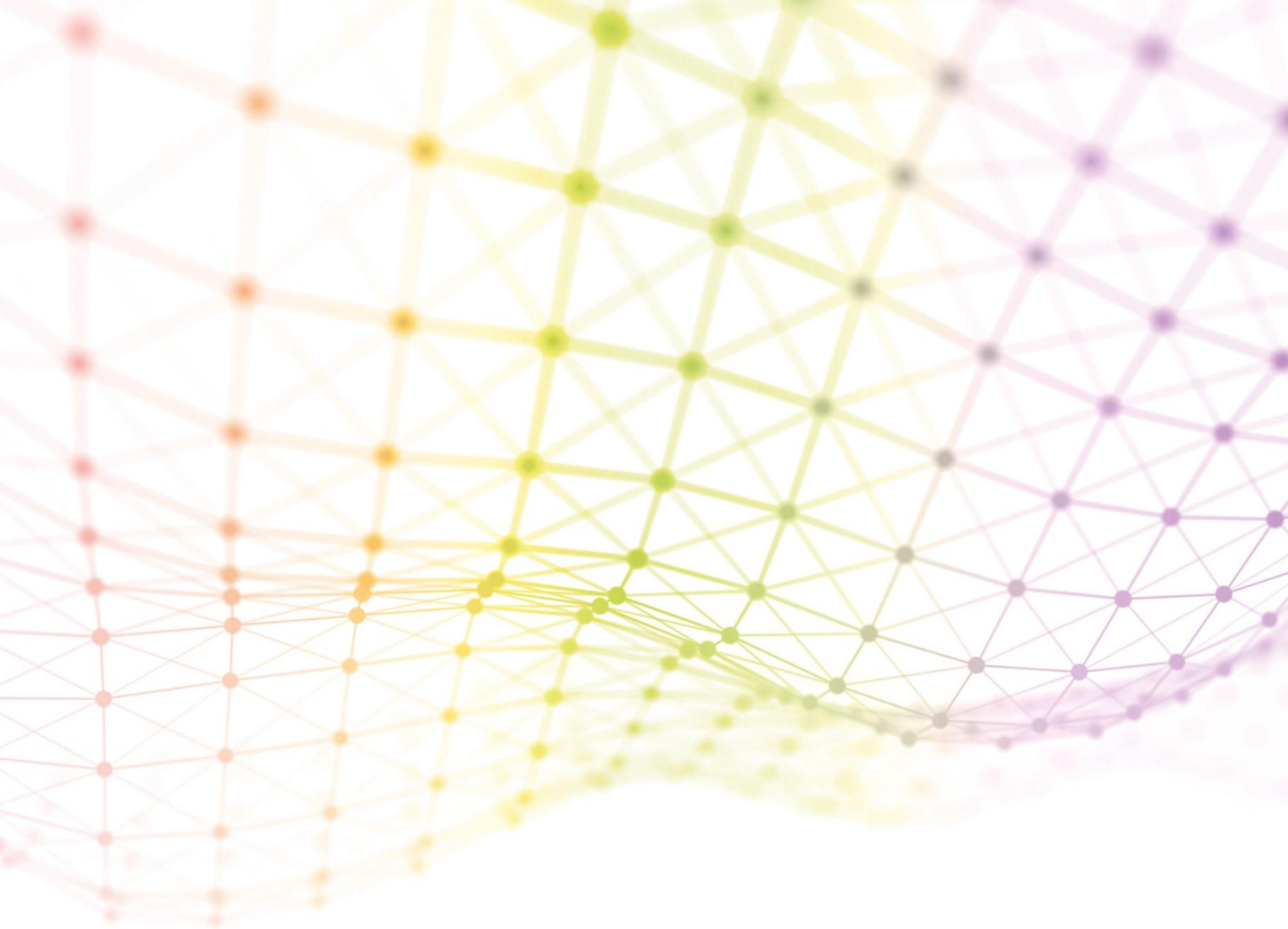
Olson, K. & Smyth, J. (2017) Within-household selection in mail surveys: Explicit questions are better than cover letter instructions, *Public Opinion Quarterly* 81(3): 688-713.

Olson, K., Stange, M. & Smyth, J. (2014) Assessing within-household selection methods in household mail surveys, *Public Opinion Quarterly* 78(3): 656-678.

Williams, J. (2019). *Five methods of within-household sampling: does it matter which one we use?* Community Life Survey, GOV.UK. https://assets.publishing.service.gov.uk/media/5ff2f881e90e0776aa14130d/ABOS_Within_household_sampling_comparative_study_-_Jan_2019_V2.pdf

Wilson, L. & Dickinson, E. (2021) *Respondent-Centred Surveys: Stop, Listen and then Design.* Sage.

Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (2021). Conducting general social surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*, *85*(2), 623-648.



University of Essex

University of Southampton

UKRI Economic and Social Research Council

NCRM National Centre for Research Methods

Office for National Statistics

UCL

Warwick The University of Warwick

Manchester 1824 The University of Manchester

City University of London Est 1894

National Centre for Social Research

LSE The London School of Economics and Political Science

Ulster University

Unil UNIL | Université de Lausanne

Ipsos

verian Formerly Kantar Public

www.surveyfutures.net