

Community Life Survey

Investigating the feasibility of sampling all adults in the household

Joel Williams



BMRB

Background to the study

The Community Life Survey was commissioned by the Cabinet Office in summer 2012. It tracks the latest trends and developments across areas key to encouraging social action and empowering communities. It also incorporates a small number of priority measures from the *Citizenship Survey*, which ran from 2001 to 2011, so that they can continue to be tracked over time.

An important strand of the Community Life contract was to investigate alternative methods of data collection. As part of this, TNS BMRB has experimented with a random sample design that uses self-completion questionnaires to collect the data instead of face-to-face interviews¹. The 2013-14 survey period had two parallel strands: (i) the standard face-to-face interview survey (n=5,000; 1,250 per quarter), and (ii) an experimental online/postal self-completion questionnaire survey (expected n=10,000; 2,500 per quarter²). The self-completion questionnaire was adapted from the interview questionnaire and an edited postal version was available on demand by calling an order line³.

For both strands, a sample of addresses was taken from the Postcode Address File and introduction letters were sent to each address. For the interview survey, this letter alerted the householder to the visit of an interviewer and provided an opportunity to opt out of the survey in advance. For the self-completion survey, the letter provided the details required to access the online questionnaire (url, user name and passcode) as well as the telephone order line for obtaining a paper version. Following the initial letter, up to two reminder letters were sent to each sampled address.

Under the interview survey model, the interviewer lists all household residents (aged 16+), enters this into a computer script and one resident is randomly selected for interview. With the self-completion design, the selection mechanism is outlined in the advance letter. In this case, a quasi-random method was used: the selected resident should be the one with either the last or next birthday, depending on instruction. To test compliance, the birth month was collected about each resident adult as part of the questionnaire⁴. From this we can see that a significant number of respondents are not the person with the right birthday and that the sampling instructions have been ignored. This is a sub-optimal outcome for a random sample survey because the credibility of its data is derived from its sampling method. A degree of non-compliance with the sampling instructions does not automatically lead to biased estimates but the *risk* of bias is certainly greater.

¹ Due to the complexity of the *Community Life* questionnaire, the principal channel for self-completion is online but an edited paper variant has been produced for those who cannot – or choose not to – complete the questionnaire online. This paper variant is available on demand.

² This is slightly higher than originally anticipated (n=9,000 per year).

³ Both versions of the questionnaire can be found in the appendix to this volume.

⁴ A simpler method was used in the edited paper variant. We asked the respondent whether their birthday was the next one in the household. If the answer was 'yes' then he/she must have been the wrong respondent (as we were targeting the one with the 'last birthday'). If the answer was 'no' and the household contained two adults, then he/she must have been the *right* respondent. If the answer was 'no' and the household contained three or more adults, then the respondent status was unknown with an assumed equal probability of being right or wrong.

Overall, 74% of respondents had the right birthday, leaving 26% as 'wrong respondents'. However, the compliance rate (the proportion following the sampling instructions) is likely to be much lower than this because (i) no sampling is required in one-adult households, and (ii) some respondents may have the right birthday by chance even if the sampling instructions are ignored. Taking this into account, a compliance rate of only 35% would produce the distribution of 'right' and 'wrong' respondents that we see.

To overcome this, TNS BMRB suggested testing a variant in which *all* adults resident in the sampled household (up to a maximum of four) could complete a questionnaire. Essentially, this removes the problematic within-household sampling stage although it brings a new risk of individuals completing multiple questionnaires to get a larger reward (£10 is offered for each completed questionnaire).

This method was tested in Q3 of the 2013-14 survey period and the results are summarised here. This report is organised around a set of key questions about this new method:

- What was the response rate and how does it compare with the standard 'single adult' design?
- Did the incentives lead to any non-compliance?
- What does the sample profile look like compared to the standard design?, *and*
- Is the data quality comparable?

Before going on to answer each of these questions, it is worth noting that the precision of estimates based on 'all adults' data will be slightly lower than it would be under the standard single adult design. This is because some of the data will be clustered by household. Individuals within the same household are more similar to each other than they are to others within the population and this similarity (or, more formally, 'non-independence') reduces the statistical value of the sample. The degree to which the statistical value is reduced will vary from measure to measure. None of this implies any additional systematic error –or 'bias' – but household clustering can introduce this as well if individuals within the household influence each other's responses or, indeed, their decision with regard to participation. This is not something that can be directly detected from the data and remains a risk, albeit one that other household surveys (e.g. the Labour Force Survey, the Health Survey for England and the ESRC Understanding Society longitudinal study) accept in exchange for lower costs and greater understanding of household dynamics. In this case, of course, TNS BMRB recommends it for a different reason: to ensure that the sampling mechanism is fit for purpose.

What was the response rate and how does it compare with the standard design?

A stratified random sample of 1,400 addresses was drawn from across England in September 2013. At least one questionnaire was completed in 341 households. This represents a 24.4% address conversion rate (95% confidence interval = 22.1-26.7%). The conversion rate is the proportion of addresses from which at least one questionnaire was obtained. For comparison, the address conversion rate for the standard design (one sampled adult per household) was 23.8% over the same period, not significantly different from the 24.4% obtained for the 'all adults' design in Q3.

Given that 90-93% of addresses are found to be residential in interview surveys, we can estimate the **household response rate in Q3 at around 26-27%**.

In total, 520 questionnaires were completed (490 (94%) online, 30 (6%) by post⁵), an average of 1.52 per cooperating household. All but two cooperating households provided information about the total number of resident adults⁶. The average was 2.07 adults per household, or 701 resident adults across all cooperating households. The limit of four questionnaires per household reduces the maximum possible number of completed questionnaires within cooperating households to 697 and means an effective **within-household response rate of 74.6%**.

As expected, the probability of completing the questionnaire was highly clustered. If one adult completed the questionnaire, the probability of other adults in the household completing it was 50% each.

The person level response rate (incorporating both the household and within-household response rates) is not straightforward to compute. Our approach is to use the comparable face-to-face interview data to estimate the number of adults in sampled addresses and divide this by the number of completed questionnaires. This leads to a **person level response rate of 21.8%**, slightly higher than the simple product of the household and within-household response rates (19.6%).

For the standard design (one sampled adult per household), the person level response rate is equal to the household response rate (i.e. 27-28%) and therefore higher than the person level response rate recorded for the 'all adults' design (22%). However, it is estimated that the standard design sampling instructions⁷ are followed in only one third of cases so a person level response rate of 27-28% is an overestimate.

In summary, we estimate that the response rate for the 'all adults' design is very similar to that of the standard design.

⁵ The equivalent Q3 postal share for the standard design was slightly higher: 8%.

⁶ There is the occasional discrepancy between questionnaires from the same household which may be inadvertent or deliberate (6/137 households with multiple returns have different totals reported). For simplicity, the highest total has been taken for this analysis.

⁷ In half of sampled households, the selected individual is the adult with the *most recent* birthday; in the other half, the selected individual is the adult with the *next* birthday.

Did the incentives lead to any non-compliance?

Because a conditional incentive was offered for each completed questionnaire, there was a motivation for some households to over-claim the number of resident adults and complete multiple questionnaires. Without a direct validation of household structures (e.g. through an interviewer visit), it is impossible to conclusively judge the extent of this non-compliance.

However, if such non-compliance were widespread, we should see a peak in the person-level response rate for four-adult households. We estimate that the overall person-level response rate was 21.8% and that this varied from 19-22% for all household sizes *bar four adult households* where the estimated person level response rate was 33% (95CI = 27-39%) *if the reported household structure data was genuine*. This may be regarded as implausibly high, providing some support for the non-compliance hypothesis. Of course, it could be argued that the 'all adults' design is simply particularly attractive to genuine four-adult households who may pool the £40 of vouchers but this does not explain why there is no effect whatsoever for three-adult households which could obtain a still substantial £30 of vouchers.

In summary, **it is likely that in a small number of cases** one adult in the household completed multiple questionnaires rather than just one for additional incentives **but not enough to represent an existential threat to the 'all adults' design**. Furthermore, with an estimated sampling instructions compliance rate of only one in three in multi-adult households, the level of 'non-compliance' is likely to be greater with the standard design.

One possible way forward is to include a line in the letter saying that 'a proportion' of the completed questionnaires will be validated through a short interviewer call or visit. **A small number of validation calls/visits would need to be paid for but that might be good value as a deterrent**. These validations should include all households in which four questionnaires have been returned.

What does the sample profile look like compared to the standard design?

A simple comparison of the gender/age profile with the standard design (after weighting the latter to compensate for differential within-household sampling probabilities) shows that **the 'all adults' design produced a more balanced sample**. Across twelve gender and age categories (six for each gender), the summed difference with the mid-2013 population distribution was 16.7 percentage points, an average of 1.5 percentage points per category. For comparison, the standard design (Q1-4) summed difference from the mid-2013 population distribution was 29.8 percentage points (an average of 2.7 percentage points per category). In fact, the 'all adults' gender/age profile had a similar error level to the interview sample, albeit with a different pattern⁸.

The proportion aged 16-24 (8.5%) is particularly better than the standard design proportion (6.4%) and closer to the interview proportion (10.9%). However, all are short of the true population proportion of 14.3%.

Table 1: Gender and age profile comparison

	All adults design (Q3) n=518	Single adult design (Q1-4) n=10,150	Interview design (Q1-4) n=5,105	ONS 2013 mid-year population totals
M 16-24	3.9%	2.3%	5.8%	7.3%
M 25-34	7.3%	4.9%	5.9%	8.4%
M 35-49	12.0%	11.2%	11.2%	12.6%
M 50-64	12.9%	15.4%	11.8%	10.9%
M 65-74	6.9%	9.5%	6.4%	5.5%
M 75+	4.2%	5.0%	4.5%	4.0%
F 16-24	4.6%	4.0%	5.1%	7.0%
F 25-34	10.2%	7.4%	8.2%	8.5%
F 35-49	14.3%	14.9%	15.0%	12.8%
F 50-64	12.5%	14.8%	12.8%	11.2%
F 65-74	6.9%	7.0%	7.7%	6.0%
F 75+	4.1%	3.5%	5.6%	5.8%

It is worth bearing in mind that the sample size for the 'all adults' data is only c500 and that a larger sample might produce quite a different profile. Gender and age are also not the only characteristics that could be assessed. However, they tend to be strong guides to the overall balance of the sample.

It is also possible to compare the results obtained from the 'all adults' design with those from the standard design (and the interview design if desired). However, the very small achieved sample size means that even quite substantial differences may be due to random sampling error rather than systematic variation. TNS BMRB recommended leaving this until a larger sample is accumulated.

⁸ The interview (Q1-4) gender/age profile summed difference from the mid-2013 population distribution was 14.3 percentage points (an average of 1.3 percentage points per category).

Is the data quality comparable to that of the standard self-completion design?

Data quality must be assessed through multiple metrics to make a rounded judgment. The following analyses are included here:

- An assessment of the consistency of household data between questionnaires submitted from the same household;
- A comparison (with the standard self-completion design) of the number of dropouts while completing the online questionnaire;
- A comparison (with the standard self-completion design) of the proportion completing the questionnaire quickly (less than thirty minutes);
- A comparison (with both the standard design and the interview design) of missing data rates;
- A comparison (with both the standard design and the interview design) of the number of responses selected from multiple response answer lists

Consistency of household data between questionnaires submitted from the same household

In 137 households, more than one completed questionnaire was received. This gives an opportunity to assess consistency of household-level data. In six cases (4% of the total) the reported number of resident adults did not match. In two cases the number of children did not match.

A tougher test is to sum the reported ages of all adults in the households and check for consistency again. In 95 out of 130 multi-questionnaire households with this data (73%)⁹, the sum of ages matched exactly. In another 26 households, the disparity was only a year or two, leaving only 9 households (7%) with larger variations.

Although household data is only a small part of the questionnaire, this **reasonable level of consistency is a good sign of data quality more generally.**

Number of dropouts while completing the online questionnaire

There was no significant difference in dropout rate between the 'all adults' design and the standard design. According to the paradata, 538 individuals started the 'all adults' online questionnaire and 490 completed it (a 91% completion rate). For comparison, 2,679 individuals started the Q2 standard online questionnaire and 2,380 completed it (an 89% completion rate). **The 'all adults' design raises no additional concerns regarding the dropout rate.**

Proportion taking less than thirty minutes to complete the questionnaire

In total, 41% of the online questionnaires were completed in less than thirty minutes, 50% were completed in 30-60 minutes, and 9% were completed in more than sixty minutes. The trimmed mean (outliers removed) was 38 minutes. A similar distribution was observed regardless of whether one, two or three questionnaires were returned from the household. However, in *households where four questionnaires were returned*, 63% completed the questionnaire in less

⁹ Note that the ages of individuals in the household was not collected from the postal questionnaire.

than thirty minutes and only 2% took more than sixty minutes. The trimmed mean was 28 minutes, ten less than the overall average of 38.

This considerable increase in completion speed is suspicious. It is plausible that if a household returns four online questionnaires its members may be particularly competent online. However, the same could be argued about a household returning three online questionnaires and yet we do not see significantly higher completion speeds in these households. The questionnaire is not strongly filtered so differences in experience that may vary systematically with household structure are unlikely to play a part here.

As with the analysis of person-level response rates, **this evidence seems to point to a small degree of non-compliance when four questionnaires are returned from a single household.** The value of the validation system described above is enhanced by this finding.

Missing data rates

Overall, 55% of completed questionnaires contained no missing data¹⁰ at all and 80% contained no more than two missing data items. A small proportion of questionnaires were more problematic: 12% contained at least ten missing data items and 4% contained at least twenty missing data items. There were one or two outlier questionnaires that would ordinarily be edited from the dataset (the highest recorded number of missing items was 130).

The missing data rate is independent of the number of questionnaires completed in the household. This may be considered evidence against non-compliance since one might expect greater use of the missing data codes if respondents are completing the questionnaires as quickly as possible. However, the structure of the online questionnaire - in which missing data codes only appear as an option if the respondent tries to move on without giving an answer - may mean speeding responders use different methods, e.g. overusing codes towards the top of a list, something that is assessed in the next section.

Number of responses selected at multiple response questions

There were 12 questions that allowed multiple responses and are present in both the online and postal versions of the questionnaire: <civpart>, <civconsult>, <fgroup>, <funpd>, <volben>, <vbarr>, <lhlp>, <ggroup>, <teuse2>, <civact1>, <civact2> and <locinvnew>¹¹.

Across these twelve questions, the mean number of responses selected was 11.9 (95CI = 11.6-12.2). If only one questionnaire was returned from a household, this mean was slightly higher than if multiple questionnaires had been returned (12.5 compared to 11.5). Although the sample size is small (n=56 across 14 households), the mean was only 9.5 (95CI = 8.6-10.4) for respondents in 'four completed questionnaire' households. This may indicate some speeding through the questionnaire (and possibly non-compliance) although it could just as easily reflect natural differences in behaviour due to differences in age profile etc. In any event, **there is no evidence of a serious additional problem with multiple response items under an 'all adults' design.**

Using the same set of questions¹², we can assess the relative likelihood of selecting the first item. A higher likelihood in multiple questionnaire households *might* reflect less attention paid to the

¹⁰ Missing data includes all uses of 'don't know' or 'refused' codes plus cases where no answer has been given to a paper questionnaire item.

¹¹ <civpart> = list of civic participation activities, <civconsult> = list of civic consultation activities, <fgroup> = list of types of group or club, <funpd> = list of unpaid help activities with groups/clubs, <volben> = list of reasons for starting volunteering, <vbarr> = list of barriers to volunteering, <lhlp> = list of unpaid help activities with individuals, <ggroup> = list of charity donation types, <teuse2> = list of things that would encourage respondent to donate, <civact1> = list of civic individual roles, <civact2> = list of civic group roles, <locinvnew> = list of organising activities.

questions although it may also reflect genuine sample differences as well. In 17% of relevant multi-code items (relevant = the question was asked), the first code is one of those selected (95CI = 16-18%). This rate hardly varies based on the number of questionnaires completed per household (15-19%), suggesting **no additional primacy effect within multiple questionnaire households.**

¹² Excluding <fgroup> where the response list was randomly ordered for the online survey.

Summary and conclusions

Overall, the evidence suggests that the 'all adults' design leads to a low level of non-compliance with the sampling instructions (as evidenced by the implausibly high person-level response rates in four-adult households and the much faster questionnaire completion). However, the limited data quality analysis did not reveal any obvious additional problems that would rule out this method.

Consequently, it seems likely that the 'all adults' design is a compatible alternative to the standard 'single adult' design. It may even yield a more balanced sample if the gender/age profile revealed by this experiment is at all typical and, of course, there is no scope for within-household sampling error as there is with the standard design.

It is the view of TNS BMRB that, if the self-completion survey is to continue, it would be better to use the 'all adults' design than the standard 'single adult' design. It effectively overcomes the problem of non-random within-household sampling without bringing substantial additional drawbacks.