Focusing on Local Areas to Address Labour Force Survey Response

1. Introduction and background

The 2014 National Statistics Quality Review (NSQR)[1] of the UK Labour Force Survey (LFS) concluded that the LFS currently enables the production of good quality estimates from the survey outputs. It also reported findings from the 2011 Census Non-Response Link Study[2] that, despite increasing non-response to the LFS, levels of bias in 2011 were actually quite low.

However, the NSQR report cautioned that, if response rates were to continue to decline, this would raise concern for both the accuracy and the precision of the survey estimates. It recommended that the LFS would benefit from an ongoing research team actively supporting continuous improvement of the LFS through dedicated research aimed at highlighting areas of concern and investigating opportunities for development.

In May 2014, a small team was established with a particular focus on addressing survey non-response differently from how it has previously been managed. The usual approach is to look at non-response at a very aggregate level. Instead, the team was tasked with looking at what has happened to response in very small/localised Interviewer Areas[3] where LFS response has fallen consistently or sharply since the 2011 census..

The approach adopted was as follows. Interviewer Areas where there has been a significant fall in response since 2011 are identified using a series of statistical filters; this ensures that the fall in response is unlikely to be due to chance. Those Interviewer Areas which have failed the filters are selected for an analysis of factors affecting the fall in their response, with the aim of understanding what has changed since 2011. On the basis of this analysis, the Field Operations Manager[4] for the Interviewer Area draws up an action plan for addressing falling response and reports regularly on progress against that action plan.

Where response has fallen in an area because practices have changed for the worse since 2011, the aim is to try as far as possible to return the area to 2011 conditions, for example by re-adopting the calling patterns or the resources they were using in 2011. Evaluation of the success of action plans includes re-running some of the initial analyses to see whether changes between 2011 and now have successfully been reversed. The project team then consider whether response rates have improved in these areas using the 2011 response rate as a target and if not whether alternative (more radical) solutions are necessary.

The three possible success criteria (for each Interviewer Area) that were set out at the start of the project are:

- the LFS wave 1 response rate increases to 2011 levels within a six month period, or

[1] http://www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-reviews/list-of-current-national-statistics-quality-reviews/nsqr-series--2--report-no--1/index.html

[2] http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/articles-and-reports/non-response-weights-for-the-uk-labour-force-survey.pdf

[3] There are 208 LFS Interviewer Areas in Great Britain. An LFS Interviewer Area is an area originally drawn to be covered by one or two field interviewers.

[4] Approximately 620 field interviewers are employed by ONS across Great Britain. There are nine Field Operations Managers across the country, each of whom manages approximately six Interviewer Managers. Each Interviewer Manager in turn is responsible for the performance of around twelve field interviewers.

- there is improvement in the LFS wave 1 response rate during a six month period that can be expected to reach 2011 levels with continuation of the actions, or
- there is no demonstrable improvement in response rate for the area during the six month period and there is therefore evidence that different approaches are required.

2. Identifying Interviewer Areas using filtering tests

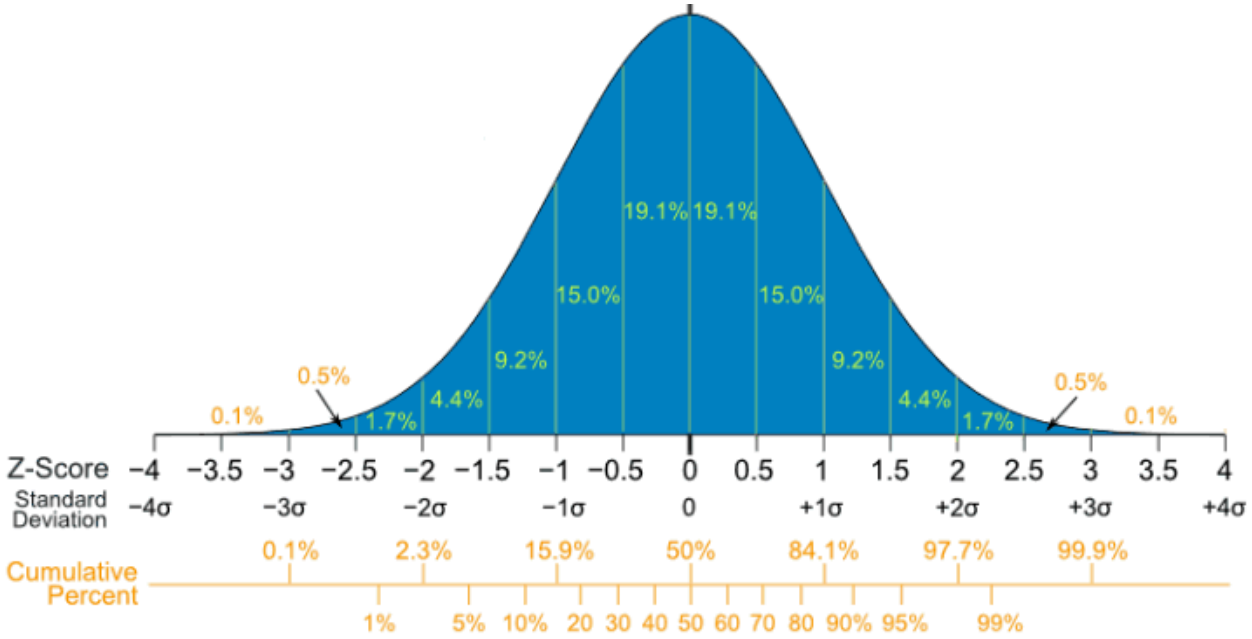
Sampled response rates in an Interviewer Area fluctuate from one period to another so may be observed to fall even where there is no 'real' decline in its underlying response rate. We want to identify Interviewer Areas where there is a 'real' fall in the response rate compared with 2011, without unnecessarily identifying those where the fall is just due to random fluctuation. Statistical filters are used to identify the Interviewer Areas we want to explore in more detail.

The overall response rate for an Interviewer Area follows a binomial distribution and can be tested using standard significance tests. Test 1 is designed to identify areas where there has been a *large* fall in response. The difference between proportions test is applied, using the binomial approximation to the normal distribution to calculate the limits.

Response for each Interviewer Area in a rolling year is compared with response in 2011. A significance level of 99%, or a Z-score of 2.33, is used for Test 1 (one-tailed, see diagram). At this level, a fall in response in an Interviewer Area of approximately 10% is necessary to reject the null hypothesis that the underlying response rate is actually unchanged since 2011. The Test 1 level is 99% to avoid rejecting the null hypothesis incorrectly too many times: there are 208 IAs, so only two would be expected to fail a 99% test by chance in any period, whereas ten would be expected to fail a 95% test. Over time, as the test is repeated each period, the chance of a 'real' change being detected increases.

Test 1 detects *large* falls in response, but to detect *consistent* smaller falls a second test was designed. Test 2 applies the difference between proportions test, but compares a quarter with the annual proportion in 2011. The level is now set at 66% or a Z score of 0.44 (one-tailed, see diagram). Under the null hypothesis, a third of Interviewer Areas are expected to fail this test due to chance. However, after applying the difference in proportions test to each quarter, a sign test is then applied. To fail Test 2, an Interviewer Area has to fail the difference in proportions test for a run of four consecutive quarters. For an Interviewer Area where there has been no 'real' change the probability of failing the proportions test by chance is a third, but the probability of failing the sign test - with four consecutive proportion test failures - is only 1.2%. Among the 208 Interviewer Areas, only two would be expected to fail this test by chance, if there had been no 'real' drop in their response rates.

The two tests identify changes in the 'real' response rate for an Interviewer Area, either by a large drop or by a consistent smaller drop or both. Combining the two tests, only three Interviewer Areas where there was no 'real' change would be expected to fail Test 1, Test 2 or both tests in any period just by chance. This is considered an acceptable level of false positives. There will of course be some false negatives: these are Interviewer Areas where 'real' response *has* fallen since 2011 but sampled response has not failed the tests. For these, achieved response is not yet a problem, so although these may eventually fail the tests, until then they are not a concern.



Since there are more Interviewer Areas which fail both tests than we have resource to analyse, we need a way of prioritising which to investigate first. Three pieces of information are used:

- Absolute response rate in the latest rolling year: Interviewer Areas with the lowest response are the highest priority.
- Total number of times Test 1 and Test 2 have been failed: Interviewer Areas which have failed these tests the largest number of times since 2012 are the highest priority.
- Total number of times response rates in a quarter have dipped below 50%: Interviewer Areas where response rates have dipped below 50% in the largest number of quarters since 2012 are the highest priority.

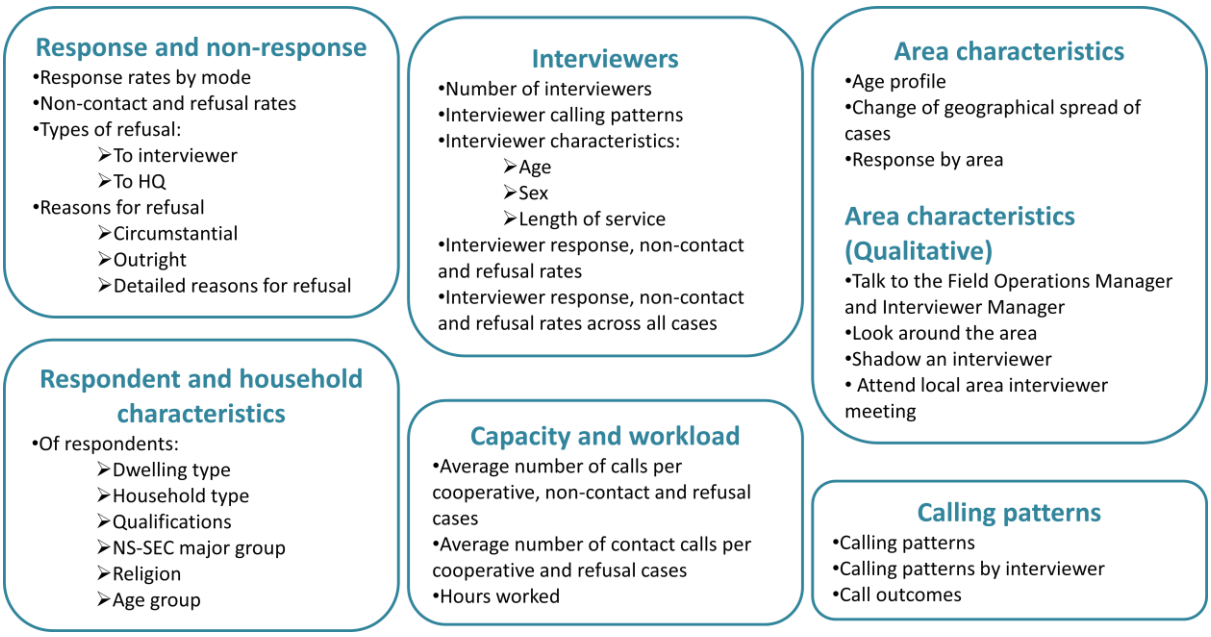
3. Analysing Interviewer Areas

Using the prioritisation described above, Interviewer Areas are selected for a detailed analysis of the factors affecting their fall in response. Analysis focuses on data collection in the latest rolling year and compares it to 2011. There are two parts to the analysis: quantitative and qualitative analysis.

Quantitative analysis uses call records, data on interviewer travel and expenses claims and LFS household data records. Each analysis explores respondent reasons for non-response (non-contact and refusal), interviewers (who they are, their performance, calling patterns and workload) and changes in respondent, household and area characteristics.

The qualitative analysis involves collaboration with the field staff and field visits. Interviewers, their managers and Field Operations Managers know the area best and can provide useful information about what they perceive to be changes in the studied area. Each analysis has also included shadowing an interviewer in the area and visiting a local interviewer meeting to gather interviewers' own views of reasons for the fall in response. Although this has been resource intensive, it gives a valuable insight into area changes and the difficulties that interviewers face in those areas.

The diagram below shows the analysis that is carried out on a selected Interviewer Area.



Once the analysis has been completed, an area report is produced by the team which summarises the analysis and what has changed since 2011 and makes some recommendations about what should happen next in that area.

4. Developing action plans

The Field Operations Manager for each Interviewer Area analysed is asked to draw up an action plan for addressing falling response in that area. In addition to actions being SMART (Specific; Measurable; Achievable; Relevant; and Time-limited), there are four additional criteria. Actions should:

- focus on setting conditions back to how they were in 2011 as much as possible (i.e. giving each address a similar amount of effort from similarly skilled interviewers)
- build on the analysis and recommendations in the area report (rather than initiating a new analysis or investigation into the reasons for falling response)
- go further than normal management practices (i.e. not *just* doing the normal observation and mentoring activities)
- allow the experiment to take place (i.e. if restoring conditions to 2011 requires temporarily moving resource from a neighbouring area, they should go ahead and show whether this works)

Where we can show that conditions have been reset as far as possible to 2011 and response still does not improve, it will indicate that there have been changes in the area which mean this approach will no longer work. In that case, we may need to adopt new measures such as targeted non-response follow-up studies, new materials or new working practices.

5. Monitoring and evaluating the success of action plans

Monitoring and evaluation is important to understand the effectiveness of the actions implemented and to make decisions about which to continue or alter. There are three parts to this:

- monthly monitoring of how actions in each area are progressing;
- three- and six-monthly reviews of the effect of the actions on key indicators; and
- an evaluation, after six months, of the success of the action plan.

Response in each of the Interviewer Areas analysed will continue to be closely monitored over the longer term.

Each month a time series of response rates in each Interviewer Area is produced and the Field Operations Managers provide progress against each of their actions. Three-monthly and six-monthly reviews provide a more detailed view of the impact and progress of the action plans. For the conditions that were highlighted initially as being a factor in the fall in response, analyses are repeated to determine whether those conditions had improved again since implementation of the plan.

Evaluation at the end of the six month period aims to understand if the changes have been successful in achieving the overall aim of the project – that is to improve LFS wave 1 response back to 2011 levels. There are two aspects to the evaluation:

- Where characteristics, conditions or metrics have changed since 2011, have the interventions and the action plan been successful in reversing the change?
- If successful, has this also improved response to (or part way to) 2011 levels?

6. Summary

The underlying premise of the project is that by setting the current conditions in each area back to the conditions in 2011, when response in these areas was better than at present, we would expect to see improved response. If reversing the changes doesn't improve response, it may indicate that the area, and people's willingness to take part in surveys in that area, has changed. We may need to adopt new approaches such as incentives targeted to the local area or increasing involvement of local managers in the day-to-day work of the interviewer. The full range of possible interventions to improve response includes more radical and expensive options such as Sunday working, significant additional field capacity and ultimately even the move to compulsory participation. The aim of this project is to be sure that we have exhausted known approaches to recovering response levels before we resort to new or more radical measures.