

Release of two methodological papers on response rate and survey bias

1. Over recent years the proportion of the public willing to take part in social surveys has gradually fallen. Traditionally, response rates have been used as a key proxy measure of survey quality – with a high response rate indicating good quality. However, previous empirical studies have suggested that using response rates as a measure of survey error or bias can be problematic. To examine and understand the relationship between response rates and survey quality, further analysis was undertaken.
2. The Scottish Government has published a methodological paper based on the impact that lower response rates would have on estimates made by the Scottish Household Survey (SHS). A similar paper on Scottish Crime and Justice Survey (SCJS) will follow.
3. Thanks to those at Ipsos MORI, outlined in the reports, who conducted the analysis and created these reports.
4. This note provides an overview and summary of the analysis presented in both reports and places the findings within a wider context.

Background

5. Data from Scottish Government household surveys feed into National Statistics publications and multiple National Performance Framework indicators. It is therefore crucial that the survey responses accurately reflect Scotland's population, as best as possible.
6. One of the main issues for survey collection is bias, due to a difference in characteristics between those who respond and those who refuse to take part. This can mean that the individuals whom you collect data from are not representative of the whole population of interest. An example of this could be systematic e.g. only collecting data at 9-5 pm on a Tuesday, would likely miss a bigger proportion of the employed population than unemployed or retired. Hence estimates made by the survey would not be fully representative. Scottish Government household surveys are designed to reduce these levels of non-response bias as much as possible by staggering visit times across the week (including weekends), visiting addresses multiple times and not using replacement addresses if an intended respondent refuses to take part.
7. Traditionally, response rates have been used as a key proxy measure of survey quality – with a high response rate indicating good quality. But the relationship between response rate and non-response bias is more complicated, with academic literature pointing to a potentially weaker relationship than previous thought. A reduced level of response only increases the potential for bias rather than directly increasing it.
8. Both of the papers, use similar analysis to explore how a response rate change of 5-10 percentage points would impact on results. This was achieved by comparing the re-weighted results based only on the sample achieved at first issue, against the final sample achieved following reissues for a range of key metrics.

9. This research provides valuable evidence for both the Scottish Government and the wider research community on what impact response rates might have on survey estimates.

Main conclusions of analysis

10. The reports show that for response rates between 5 and 10 percentage points lower, estimates made from the survey are broadly very similar. This is in terms of the absolute percentage or value estimates themselves and also as a share of normal survey error. This seems to support previous literature that there is not a simple relationship between response rates and survey bias.

11. The Scottish Household Survey and Scottish Crime and Justice Survey undertake fieldwork by issuing addresses to interviewers. Interviewers visit addresses multiple times in an attempt to achieve an interview. Normally 6 or more visits are made when addresses are first issued to make contact with a householder. Most addresses where the initial interviewer is unsuccessful at getting an interview are reissued to a second (or third) interviewer. These are commonly referred to as reissues. They include addresses where no contact was made and a proportion of refusals where a householder initially declines to take part.

12. Randomly selected probability samples such as SHS and SCJS, carry a known sampling error with each estimate. As we are only taking a sample of the population rather than collecting data from every person (e.g. a census), there is potential for error in our estimates. Due to the sampling approach in both surveys we are able to quantify this. Sampling errors are usually presented as being +/- 2 standard errors from the mean, which allow us to create confidence intervals on our estimates.

13. As well as presenting the absolute percentage differences made from the whole sample and a sample without reissue (between 5 and 10 percentage points lower), these papers also present this in respect of sampling error. Most estimates calculated with lower response rates were within 2 standard errors of our current estimates, hence lower than the sampling error normally reported.

14. Overall, this means that the estimates we would produce with response rates which are 5-10 percentage points lower to those actually achieved, are similar to those we currently present in the confirmed National Statistics. This is in part due to response rates at first issue generally being lower for men, younger age groups and those living in urban and deprived areas. As these differences are already known, they are accounted for in the survey weighting strategy.

15. Although the estimates made would be similar, the precision and statistical power of these estimates may change. A lower number of people responding (sample size) would mean that confidence intervals would be larger, hence the estimates would be less precise. We estimate the likely response rate to the survey in order to draw a sample size prior to surveying, therefore any unexpected drops in response rate will affect the precision of the estimates and our ability to make statistically valid comparisons between groups.

Interpretation

16. Response rates to large-scale national surveys have been falling gradually over the past 20 years, which means our knowledge and interpretation of non-response bias is becoming more important.

17. Acceptable levels of response rate must be balanced against two other factors; i) the level of survey error (precision) which is mainly based on sample size and, ii) the cost and value for money in visiting the level of addresses required for a certain precision.

18. Although the analysis presented suggests that a lower response rate does not affect estimates and therefore we might be able accept a lower response rate in order to gain greater precision, there a few reasons why this may not hold.

19. These studies represent a relative change in non-response for the specific estimates analysed at the current response rates (between 60 and 65%). We have no evidence to suggest that a further drop in response e.g. from 55% to 45% would mirror the differences seen in the analysis.

20. Theoretically, we may assume that people who take more effort to respond to a survey and require a reissue, are more likely to be similar to those who refuse to take part. This means by currently collecting data via reissues we are not only gaining value from increased sample size and increased response rate, but we also gain some valuable insight into the groups of people who will respond with more effort. This could be useful when understanding any changes to survey approaches in future years. We are continuing to monitor response rates and aim to maximise these within each survey, but this analysis provides welcome evidence in understanding any changes which do occur.

COVID-19

21. These analyses present results for surveys undertaken prior to the COVID-19 pandemic.

22. It is expected that the impact of COVID-19 on certain individuals willingness and ability to take part in a face-to-face household survey, will impact on response at first and subsequent issues of addresses.

23. Response rate may change when a return to face-to-face interviewing is possible. The response rates seen at first issue in previous years (~50%) may require reissues to reach the same levels when interviewing can recommence.

24. The differences in characteristics between individuals taking part at first issue and subsequent issues may also change due to COVID-19.

25. There could also be a positive effect on response rates. A new understanding of the importance of good data collection and its use for the public good, could lead to an underlying increase in willingness to take part.

26. Overall, we see these reports as providing vital additional evidence on the link between response rate, survey precision and value for money. They provide some reassurance that small changes in survey response rates year-on-year may not have

a significant impact on data quality and survey estimates, though caution should remain around making large changes to survey methodology.

Office of the Chief Statistician and Data Officer, Scottish Government
June 2020