**Appendix 1**

*Sampling design*

Given the survey is targeting such a narrow age range a quota design offers a much more practical approach to carrying out the survey than traditional pre-selected sampling, enabling fieldwork to be completed more quickly and at a lower cost. Indeed, the costs of carrying out the survey using a random design would be prohibitive. Although historically quota sampling may have yielded a biased sample, it is now argued that the a rigorously designed quota sample is as good as a probability sample (Raab, 2009).

The proposed design for the quota sample was the same as it would be for an equivalent probability sample at all stages other than the final selection and recruitment of respondents. The design was as follows:

1. all census datazones in Scotland were stratified by local authority, and allocated to the Scottish Household Survey analysis regions, comprised of large local authorities and combinations of smaller local authorities. The purpose of this stratification was to ensure that the sample is distributed proportionately across Scotland.
2. within each of these regions, datazones were selected with probability proportionate to the proportion of the aged 18 to 34 years household population. This provided us with the basic sampling points for the survey.
3. each individual datazone was allocated a unique sample point number and had quotas set to reflect the structure of the local population based on Office of National Statistics mid-year estimates. We selected quotas by sex, age (three quota groups – most likely 18-23; 24-29; 30-34), and work status (three quota groups – most likely working full time; in full time education; not working full time/unemployed).
4. The target profile and achieved sample is presented in the table below. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates>).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Quota Breakdown** | **Target****Profile** | **%** | **Achieved****(unweighted)** | **%** |
| Male | 1,773 | 50.6% | 1,740 | 49.6% |
| Female | 1,731 | 49.4% | 1,768 | 50.4% |
| 18-23 | 1,298 | 37.0% | 1,373 | 39.1% |
| 24-29 | 1,258 | 35.9% | 1,259 | 35.9% |
| 30-34 | 948 | 27.1% | 876 | 25.0% |
| Working FT | 1,808 | 51.6% | 1,739 | 49.6% |
| Student FT | 668 | 19.1% | 609 | 17.4% |
| Not working FT | 1,028 | 29.3% | 1,160 | 33.1% |
| ***Total*** | ***3,504*** | ***100%*** | ***3,508*** | ***100%*** |

1. the position of each sample point was plotted on a map to illustrate the distribution of the sample.
2. interviewers were able to approach any household within the sampling point to achieve interviews in line with the quota controls for that datazone. They were issued with a full listing of addresses within the sampling point (as opposed to a small number of addresses in each sampling point as would be done in a random pre-selected design). Interviewers’ work was tightly controlled and they had to work in the sampling point selected – this is more robust than other approaches which allow interviewers to work in broader areas (e.g. within a particular electoral ward or constituency).
3. at those addresses we instructed interviewers to achieve a target number of interviews, with individuals that fulfil specified quotas (sex, age, work status) that had been set to reflect the demographic profile of that area. In a quota survey design, interviewers can call at as many addresses as necessary at each sampling point to achieve their interviews. Interviewers are required to leave a minimum of two addresses between interviews and visit at both the weekends and evenings to give them the best chance of meeting their quotas.
4. due to the special access permissions required, the design does not include people aged 18 to 34 years living in institutions, for example prisons or university halls of residence.