SCOTTISH HOUSEHOLD SURVEY

Methodology and fieldwork outcomes 2009/2010

Steven Hope

Ipsos MORI Scotland

Ivonne Nava-Ledezma

Ipsos MORI Sampling and Statistics

A National Statistics Publication for Scotland

Contents

List of Tables

1 Survey overview

1.1 Background to the SHS

The Scottish Household Survey (SHS) is a continuous, large-scale social survey of the composition and characteristics of Scottish households, the attitudes and behaviour of adults, and to a lesser extent, of children within these households. It was first commissioned by the Scottish Executive (now the Scottish Government) in 1998 and has been collecting data and reporting since February 1999.

The specific aims of the survey are:

- to provide household and individual information previously unavailable in Scotland, particularly to support the work of the Scottish Government's transport, communities and local government policy areas and the work of the Scottish Parliament
- to permit disaggregation of such information both geographically and in terms of population sub-groups (such as families with children or the elderly)
- to allow the relationships between social variables within households to be examined, supporting cross-departmental and inter-departmental policies such as those on social justice
- to allow early detection of national trends
- to allow detailed follow-up surveys of sub-samples from the main survey sample, if required.

The survey is funded by the Scottish Government and is carried out by a consortium involving Ipsos MORI and TNS BMRB[1].

Technical Reports have been published for each year of the survey covering the survey methodology, fieldwork outcomes and the questionnaire used. This report covers the methodology and fieldwork outcomes for the survey. On the whole, the methodology for the survey remains fixed for each two-year sweep and over the life of the survey, since 1999, there have only been relatively minor changes to the methodology following refinements to the sampling assumptions. More substantial changes were introduced in 2007 and these are discussed below.

A second publication provides details of the survey questionnaire used in 2010, including details of changes between 2006 and 2010.

[1] TNS BMRB was formerly TNS System Three.

1.1.1 Reporting conventions

In tables showing percentages as whole numbers, zero values are displayed as a dash (-), values between 0% and 0.5% are displayed as 0% and values between 0.5% and 1% are rounded to 1%. Where percentages are shown to one or more decimal place, the final digit will have been rounded up or down. As a result of rounding within tables, the sum of individual items may not equal the totals for rows or columns.

2 Sample design, selection and allocation

The sampling requirements of the Scottish Household Survey (SHS) are specified in terms of providing data with a level of precision equivalent to what would be achieved by a simple random sample (SRS) of a particular size. Five requirements were identified in the survey specification and although the preference was for unclustered sampling, contractors were invited to propose design solutions that would most cost-effectively meet these requirements.

The sampling requirements were to provide a minimum of the SRS equivalents of:

- 2,500 interviews in Scotland as a whole for each quarter in each year
- 500 interviews in each local authority with over 120,000 households in each year
- 500 interviews in each local authority (regardless of size) in each two-year period
- 500 interviews in each category of the six-fold urban rural classification in each year, and
- 2,000 interviews in the 15% most deprived areas of Scotland, taken together as a group, in each year.

These needs are met by a combination of strategies involving:

- disproportionately stratifying the sample to ensure that each local authority and each other geographical area has enough interviews in each survey period
- using a combination of unclustered and clustered sampling to maximise the cost-effectiveness of fieldwork by minimising the impact of weighting and clustering on the survey's effective sample size
- modelling the impact of different sampling methods and combinations of clustered and unclustered sampling to assess the impact of the survey design on survey precision at appropriate geographical levels at each of the specified points in time i.e. nationally each quarter, for the large local authorities each year etc.
- allocating the selected sample appropriately over survey periods to provide the appropriate level of clustering at each period.

The survey has a number of other requirements that have been retained from previous sweeps such as:

- the sample should be fully national in character – while it is not uncommon for national surveys to exclude the area north of the Caledonian Canal or to be restricted to the Scottish mainland and the larger inhabited islands, the SHS includes all parts of Scotland, including small inhabited islands.
- the sample should be capable of producing data which are representative both of Scottish households and the adult (aged 16+) population resident in private households.

Each of these features of the sampling is discussed more fully below.

2.1 Stratification by local authority

In general, stratifying a sample by some known variable should improve the precision of survey estimates because structuring a sample in this way can be no worse than would be achieved by a random allocation. For example, selecting a national random sample **should** result in a geographical distribution that reflects the distribution of population but it may not and some random deviation from the known distribution would be expected. Stratification removes this element of chance by assigning sample to geographical areas and drawing smaller samples within each of the strata. In this way the sample distribution must reflect the known population distribution.

The sampling requirements for the SHS to provide at least the SRS equivalent of 500 interviews in each local authority indicates a need for sample to be allocated between local authorities in a way that does not reflect the distribution of the population. For a given sample size, meeting this need requires smaller local authorities to have more interviews than a proportionate allocation would give them and, as a result, larger local authorities have fewer interviews.

Analysis at a national level requires the data to be weighted so disproportionate stratification reduces the precision of survey estimates, making the gross sample equivalent to a smaller simple random sample. This impact needs to be considered in the meeting the survey's sampling requirements.

2.1.1 Disproportionate stratification between local authorities

The underlying principle here is that the allocation of interviews by local authority area should be broadly proportionate to the number of households, except where the resulting sub-sample in any particular area would fall below a pre-determined accuracy threshold (the equivalent of a simple random sample of 500 interviews). The allocation was carried out in the following way.

1. The sample design needs to take account of the allocation of interviews to each local authority and for each allocation, the resulting impact on the number of interviews derived from the smallest category of the urban rural classification and the 15% most deprived area and the number of interviews achieved each year in the five largest local authorities.

2. These estimates need to take account of the impact of clustering and weighting at each level and at each relevant time period to ensure that the sample will, after accounting for weighting and clustering, meet the SRS equivalent sample size requirements specified.

3. Having established that a particular design will meet the sampling objectives at each geographical level and at each time period, the total sample is built from the combinations of individual local authority samples.

Table 2-1 shows the expected distribution of sample by local authority at the end of each two-year sampling period.

Table 2-1: Projected two-year achieved sample size by local authority and SRS equivalent sample over target periods

	Total households	% of ints from unclustered sample	Sample size Gross in target period	Sample size SRS equivalent in target period
Large local authorities – minimum of 500 SRS equivalent interviews each year				
Edinburgh, City of	209,502	100%	1,129	1,129
Fife	153,040	63%	914	825
Glasgow City	276,291	100%	1,489	1,489
North Lanarkshire	134,700	100%	726	726
South Lanarkshire	128,238	100%	691	691
Other local authorities – minimum of 500 SRS equivalent interviews after two years				
Aberdeen City	98,859	100%	1,065	1,065
Aberdeenshire	92,067	0%	1,344	992
Angus	47,861	63%	572	516
Argyll & Bute	41,864	0%	678	500
Clackmannanshire	20,876	100%	500	500
Dumfries & Galloway	65,487	30%	866	706
Dundee City	67,032	100%	722	722
East Ayrshire	51,345	37%	662	553
East Dunbartonshire	42,763	100%	500	500
East Lothian	38,757	25%	622	500
East Renfrewshire	35,388	100%	500	500
Eilean Siar	11,360	100%	500	500
Falkirk	63,684	100%	686	686
Highland	92,514	22%	1,255	997
Inverclyde	37,883	100%	500	500
Midlothian	33,229	66%	549	500
Moray	36,515	25%	622	500
North Ayrshire	60,027	70%	702	647
Orkney Islands	8,380	100%	500	500
Perth & Kinross	60,866	36%	788	656
Renfrewshire	75,867	100%	818	818
Scottish Borders	48,790	28%	648	526
Shetland Islands	9,287	100%	500	500
South Ayrshire	50,754	69%	595	547
Stirling	37,321	54%	568	500
West Dunbartonshire	41,112	100%	500	500
West Lothian	65,030	70%	761	701

	Target	Target period	Gross sample	SRS equivalent
National	2,500	Quarterly	3,552	3,052
Smallest category of urban rural classification	500	Annually	1,301	650
15% most deprived datazones	2,000	Annually	2,101	1,881

2.1.2 Clustered and unclustered sampling

In previous years, the SHS has used a combination of clustered and unclustered sampling but the tendering for the new contract allowed some reconsideration of the approach used. Although the survey

specification expressed a preference for a wholly unclustered sample, it was necessary to test a number of designs to identify the extent to which a more cost-effective design could be achieved by retaining an element of clustering in areas where this would produce fieldwork efficiencies.

An unclustered sample involves, within each of the primary strata (the local authorities), either selecting a simple random sample of addresses (n) from all possible addresses (N) or selecting a systematic sample by identifying a random starting address and selecting addresses with a fixed interval equal to N/n. Systematic sampling has the advantage that by ordering the addresses by some characteristics (such as deprivation or postcode) it is possible to achieve a further level of implicit stratification as the ordering and fixed interval sampling ensures that the variables used to order the list are represented in proportion to their prevalence in the population.

Clustered sampling is a two-stage process of selecting a sample of geographical units within which a sample of individual addresses is selected. The benefit of clustered sampling is to reduce the mean distance between sampled addresses, improving the efficiency of survey fieldwork by reducing interviewer travel time between addresses and increasing the number of times interviewers can try to make contact at sampled addresses. Although clustering offers an administrative benefit, this comes with a statistical penalty because clustering increases the likelihood that the achieved sample will be more variable than the total population. The reason for this is that the number of sampled clusters is generally small relative to the total number of potential clusters and, within each cluster, the sampled addresses are likely to be more similar to each other than they are to the addresses sampled in other clusters. This increased variability can be estimated and compared with the level of variability of a simple random sample to give an estimate of the achieved sample's simple random sample equivalence.

The extent to which a sample should be clustered therefore requires some comparison of the precision of the sample and the cost of achieving that level of precision – a clustered sample is only more cost-effective than a simple random sample if it can achieve the same level of precision as a simple random sample at lower cost. More generally, the most cost-effective design is the one with the lowest cost per effective interview.

There are two key variables that influence whether a sample should be clustered.

- Population density – the extent to which the population is naturally concentrated in a geographical area or spread across a number of small settlements or individual houses separated by large distances.
- The sampling interval – the size of the survey sample in relation to the population from which it is selected because for a given population, a larger sample will result in sampled addresses being closer together, reducing the administrative gains from clustering compared with an SRS.

The implication of this is that there is likely to be no efficiency gain from clustering a large sample in an urban area whereas there is greater likelihood of efficiency gains from clustering a smaller sample in a dispersed or rural area.

This broad approach was used in the SHS sampling using the Scottish Government Urban Rural Classification[2] to identify areas where sample should be clustered or unclustered. The general approach was that areas classified as 'large urban areas' or 'other urban areas' would use unclustered sampling while areas in the other four categories (accessible small towns, remote small towns, accessible rural and remote rural) would use clustered sampling.

This was applied within local authorities meaning that the sample was further stratified within each local authority using the urban rural classification and that each local authority potentially contains a combination of clustered and unclustered sampling.

In practice, this general approach is modified in two ways:

- Where more than 80% of households in a local authority fall into the 'urban' or 'non-urban' category, the whole local authority is treated as that category
- The three island local authorities (Eilean Siar, Orkney Islands and Shetland Islands) and Clackmannanshire use wholly unclustered sampling even though their urban rural classification suggests that they should use wholly clustered sampling. In these areas, the sampling interval is between 1 and 6 households and 1 in 8 households, which means that clustered sampling would be no more efficient than unclustered sampling. Indeed, every available cluster would be selected.

2.2 Allocating sample to different time periods

The consideration of clustering is complicated for the SHS because the sampling requirements are expressed in terms of the equivalent of a simple random sample at different points in time. Consideration needs to be given to the structure of the sample at these time points and the extent of clustering in the sample taken into account. For example, although the sample in Glasgow might be selected without clustering, in practice, the two-year sample is allocated to survey years, 'batched' into interviewer allocations and these are then assigned to months of the year – creating clusters of addresses. Thus, each month, the sample in Glasgow is made up of clusters of addresses, with the accumulated sample becoming progressively less clustered throughout the year. In practice then, the sample in Glasgow is only completely unclustered after a full year. Even if the samples in all local authorities were sampled without clustering, the quarterly samples would be clustered and this needs to be considered in terms of the ability of the design to meet the quarterly target of an SRS equivalent of 2,500 interviews.

The way in which sample is grouped into clusters optimises the extent of unclustered sampling in appropriate areas to coincide with reporting requirements.

- Large local authorities identified as requiring separate reporting of results each year – with the exception of Fife, the sample is derived from wholly unclustered sampling. All of the unclustered sample is first randomly allocated to years, grouped into the most efficient fieldwork batches and then

[2] http://www.scotland.gov.uk/Topics/Statistics/About/Methodology/UrbanRuralClassification

these batches are allocated to months within each year. In Fife, the clustered sample is treated in the same way as in all other local authorities.

- Other local authorities, which require separate reporting after two years – addresses in datazones classified for unclustered sampling are combined, sorted by deprivation indicator and a systematic sample selected. These addresses are batched, batches allocated to survey years and then to months within each year. Within each local authority, all of the datazones classified for clustered sampling are grouped and a sample of datazones selected with probability proportionate to size. Within the sampled datazones, a systematic sample of addresses is selected. Sampled datazones are randomly allocated to survey years and then to months within each year.

This has implications for how much of the SHS sample is unclustered at any point in time. In each quarter, the whole sample is clustered. Each year, only the sample from four of the five large local authorities is unclustered. In the unclustered parts of other local authorities, the addresses are first assigned to batches and then the batches are allocated between years, creating a clustered sample. After two years, the five large local authorities and the unclustered samples from all other local authorities are unclustered, leaving only the sampled datazones as clustered samples. After one year approximately 34% of the achieved sample should be from areas of unclustered sampling but after two years this will increase to over 70%.

2.2.1 Allocating sample across the calendar year

As the fieldwork for the survey runs throughout the calendar year, it is important to ensure an even distribution of batches over time and to ensure that the allocation of batches is geographically and demographically representative. There are two main reasons for this: an uneven distribution would jeopardise the requirement for the sample to be representative of the national population each quarter and some of the variables measured by the survey are likely to exhibit seasonal patterns – e.g. rates of economic activity, modes of transport.

The procedure for allocating primary sampling units (PSUs) to months of the year is derived from that developed by the Office for National Statistics (ONS) in managing the Family Expenditure Survey (FES)[3] and differs only in the need for the SHS sample to be spread evenly across 24 rather than 12 months.

Batches of addresses are allocated to survey years and within each year, sorted by local authority and by deprivation within each authority. The list of batches is then labelled with a random permutation of the numbers 1 to 12 representing the twelve months covered by the fieldwork. This permutation is generated with certain properties to avoid 'bunching' of interviews within particular quarters:

- the first four months are from different quarters
- every subsequent month is from the same quarter as the one four places before.

[3] See C. Lound, 'Allocating primary sampling units for annual surveys to months of the year', *Survey Methodology Bulletin*, No. 39, July 1996.

The example given by ONS (and used to allocate the 1996 FES) is as follows:

Table 2-2: Procedure for allocating PSUs by month of fieldwork

Position in list	Month	Quarter
1, 13, 25, etc.	10	4
2, 14, 26, etc.	8	3
3, 15, 27, etc.	5	2
4, 16, 28, etc.	1	1
5, 17, 29, etc.	11	4
6, 18, 30, etc.	7	3
7, 19, 31, etc.	4	2
8, 20, 32, etc.	2	1
9, 21, 33, etc.	12	4
10, 22, 34, etc.	9	3
11, 23, 35, etc.	6	2
12, 24, 36, etc.	3	1

As this sequence can be added automatically to the sampling procedures for the survey, no time is spent manually assigning batches to particular months.

2.3 Allocating sample between contractor organisations and questionnaire modules

Once all of the sampled addresses are batched for fieldwork, the batches are randomly assigned to the contractor organisations in proportion to each contractor's fieldwork commitment (in 2009/2010 this as 40% of interviews by TNS BMRB and 60% by Ipsos MORI).

The contract for 2007-2010 envisaged a requirement for greater modularisation of the SHS than had previously been the case. Modularisation requires the ability to only ask questions of a random sub-sample of respondents and for those sub-samples to be based on time periods or nationally representative sub-samples. One need identified in the survey specification was to create a module to measure participation in culture and sporting activities.

To meet the need for modularisation, all sampled addresses were randomly assigned to one of 10 sub-samples or interview streams, which could be used as the basis for assigning samples of respondents to particular blocks of questions. For example, the Culture and Sport module is intended to provide representative data on adults' participation and this is achieved by assigning the module to streams 1 and 6 meaning that a random sub-sample of 1 in 5 addresses and (assuming no difference in response rates) 1 in 5 interviews will be directed through those questions.

Other smaller blocks of questions are asked of sub-samples at various points in the questionnaire and the published version of the questionnaire indicates where and at what points in time streaming is used.

2.4 Sampling from the Postcode Address File

The Small User File of the Postcode Address File (PAF) is the most widely used sampling frame for general population surveys.[4] The principal advantages of the PAF are its completeness (it is estimated to miss the addresses of only 2% of the adult population and is updated every three months) and lack of bias (those addresses which are missing from the PAF are not as likely to be concentrated among particular types of people). There are, however, a number of issues arising from its use.

2.4.1 Deadwood

The Small User File of the PAF, which forms the basis of the sample of addresses, contains a number of addresses that are not residential (usually small shops and offices) or which have been demolished or are unoccupied. In addition to PAF addresses that are out of scope for any household survey, there are also addresses that are deemed out of scope for this survey. These are mainly second homes or holiday rental homes. In total, the extent of 'deadwood' in the PAF varies by area, but is usually estimated at around 10% in national samples. This is accounted for by drawing slightly more addresses than the response rate target would suggest. Thus, if the response rate target is 70% and deadwood is estimated to be 10% then for every 100 interviews to be achieved, 160 addresses are issued to interviewers rather than the 140 suggested by a response rate of 70% alone.

In practice, the number of additional addresses selected to allow for deadwood varies by local authority based on the contractors' experience of SHS fieldwork carried out between 1999 and 2005 (the most recent data available when the sampling for the 2007-2010 period was specified).

2.4.2 Accuracy and completeness

The sample for the survey is drawn for each two-year fieldwork period and so may exclude households in newly-built housing entering the PAF during the period of the survey. However, data suggests that new housing accounts for only around 1% of the housing stock in any year[5] and the impact of this is reduced by the fact that new properties are often entered onto the PAF some time before they are actually completed.

2.4.3 Exclusions

Samples of the general population exclude prisons, hospitals and military bases. While prisons and hospitals do not generally have significant numbers of private households, the same may not be true of military bases. These are classified as Special EDs in the Census and account for just 0.5% of the population. Interviewing on military bases would pose fieldwork problems relating to access and security so they are removed from the PAF before sampling.

---

4 Small users are postal addresses that receive fewer than 20 items of mail per day.

5 http://www.scotland.gov.uk/Topics/Statistics/Browse/Housing-Regeneration/HSfS shows new housing completions of 22,019 in 2008 and a stock of dwellings of 2,430,000 at the end of 2007, making new housing 0.96% of the total 2007 stock.

Specific accommodation types — The following types of accommodation are excluded from the survey if they are not listed on the Small User file of the PAF:

- nurses' homes
- student halls of residence
- other communal establishments (e.g. hostels for the homeless and old people's homes)
- mobile homes
- sites for travelling people.

Households in these types of accommodation are *included* in the survey if they are listed on the Small User file of the PAF and the accommodation represents the sole or main residence of the individuals concerned. People living in bed and breakfast accommodation are similarly included if the accommodation is listed on PAF and represents the sole or main residence of those living there.

Students' term-time addresses are taken as their main residence (in order that they are counted by where they spend most of the year). Since halls of residence are generally excluded, however, there will be some under-representation of students.

2.5 Multiple dwellings

There are potential problems associated with the fact that a single entry on the PAF may actually represent multiple dwellings or that a dwelling may contain multiple households. For example, an address listed as 14 Milton Street may consist of a tenement block containing 8 separate flats. Often, the existence of these additional addresses is indicated in the PAF in a field known as the Multiple Occupancy Indicator (MOI). To ensure that such households had an equal chance of inclusion, it is necessary to weight the address when drawing the sample. Thus 14 Milton Street would appear 8 times. In the address listings issued to interviewers, such addresses appear as '14 Milton Street - 3 of 8' etc., with interviewers given clear counting procedures for identifying the relevant selected dwelling.

Where the MOI is correct, this procedure is unproblematic. Sometimes, however, the MOI is incorrect or missing (in about 2% of cases) and the true number of dwellings at an address is only discovered once the survey is in the field.

Where an interviewer finds that the MOI is different from the actual number of dwellings observed (and there is more than one dwelling) he or she contacts the office where the correct details are used to randomly select one of the dwellings.

2.6 Respondent selection

As the survey is intended to collect information both about the structure and characteristics of Scottish households *and* about the people who occupy those households, the interview has a two-part structure.

The respondent for the first part of the interview must be a householder – generally the Highest Income Householder or their spouse or partner[6]. For the second part of the interview, one adult (aged 16+) member of the household is selected at random by the CAPI script. If this person is not available at the time, the interviewer will call back to complete the interview at a later date if necessary.[7]

[6] This must be a person in whose name the accommodation is owned or rented or who is otherwise responsible for the accommodation. The Highest Income Householder (HIH) is taken as the household reference person for the first part of the interview. In households with more than one householder, the person with the highest income is taken as the household reference person. If householders have exactly the same income, the older or oldest is taken as the household reference person.

[7] The selection of the random adult is slightly more complex than this. The random adult needs to be one of the adult household members who is aged 16 years or over, is normally resident during term time (if a student) or has not been living outside of the household for 6 months or more.

3 Data collection methods and instruments

3.1 Use of Computer Aided Personal Interviewing (CAPI)

In common with many other large-scale government surveys, the SHS is carried out using Computer Aided Personal Interviewing (CAPI). This offers a number of important advantages over traditional pen-and-paper interviewing for a survey of this kind. These include the following:

- CAPI allows greater complexity in questionnaire design, since routing and 'loops' in the interview can be automated and thus effectively hidden from the interviewer. It also eliminates the need for complex selection procedures during an interview, since random selection can be built into the program.
- Overall data quality is improved because the need for a separate data entry stage is eliminated and because automatic routing and range and logic checks reduce the scope for interviewer error.
- Preliminary data are available at the end of each day's fieldwork and the lack of a separate data entry stage allows faster turnaround of results more generally.
- The CAPI system generates detailed information about the timing and duration of interviews, allowing fieldwork to be monitored more closely.

The increased complexity of the SHS that resulted from the introduction of the Culture and Sport module and the increased level of sub-sampling within other parts of the questionnaire meant that the In2itive CAPI software used since 1999 became problematic. The difficulties with In2itive are outlined below. In 2008, TNS-BMRB developed a version of the SHS script using the Nipo system used for TNS' other work and began using this at the start of 2009. In the first half of 2009, Ipsos MORI developed a version of the script using Quancept and this was implemented in July 2009.

3.2 Questionnaire development and changes

The original SHS questionnaire was developed between August and December 1998 by the contractors, working in conjunction with the Technical Group for the survey from the Scottish Executive. This followed a broad consultation exercise in which interested parties from a range of policy areas, academia, voluntary organisations and other bodies were invited to propose topics or specific questions for inclusion in the survey.

The core of the SHS questionnaire is intended to remain constant, but there is also scope for the inclusion of different modules over time. A simplified version of the questionnaire for each survey year is provided as a separate document (see www.scotland.gov.uk/SHSPublications) as well as a detailed Topic List giving a broader look at past and present SHS topics can be found on the SHS website at www.scotland.gov.uk/SHSTopics.

3.3 Questionnaire structure, length and content

It was noted earlier that the questionnaire is in two parts: the first collecting information about the composition and characteristics of the household from the Highest Income Householder or their spouse/partner; the second focusing mainly on the attitudes and experiences of a random adult member of the household. The former is intended to generate data representative of Scottish households and the latter data representative of the Scottish adult population resident in private households. It should be noted, however, that for reasons of space a handful of 'household' questions are also asked of the 'random adult'. These address household events or characteristics about which any adult member of the household would be likely to know.

A simplified version of the questionnaire can be found in the publications section of the SHS website (www.scotland.gov.uk/SHSPublications) The broad topic areas, however, are as follows.

In the first half of the interview (with the highest income householder or spouse/partner), respondents are asked about:

- household composition and characteristics of household members
- type of property/accommodation
- internet access
- recycling
- cars in household and access to public transport
- children in the household, childcare, satisfaction with schooling and travel to school
- health problems and caring responsibilities
- employment status of the highest income householder
- household income from employment and other sources
- savings and household finances, including mortgage and rent.

In the second half of the interview (with the 'random adult'), respondents are asked about their own:

- housing experiences, including homelessness
- perceptions of the local area
- perception and experiences of neighbourhood disputes
- education qualifications
- use of internet
- travel to work or education and use of private and public transport
- congestion

- travel on the previous day
- perceptions of services and local government
- volunteering
- participation in sports and cultural activities
- health problems and caring responsibilities
- employment status
- individual income from employment and other sources.

3.4 Problems and errors in the survey scripts

New versions of the survey script were introduced in 2009. Between June and December 2008 TNS-BMRB reprogrammed the SHS script using their in-house CAPI software NIPO. This was tested against the previous In2itive script and interviewers briefed in the run-up to the start of 2009 fieldwork. During the first half of 2009, Ipsos MORI reprogrammed the script using Quancept, the software developed by SPSS to replace In2itive and which had been running in parallel with In2itive across the company for the previous two years and to which all projects were being moved, phasing out In2itive. This new script was introduced in July 2009.

The introduction of the new versions of the script in NIPO and Quancept did not result in any data issues and no further developments, other that minor question revisions, took place in 2010.

3.5 Survey fieldwork

The main fieldwork for the survey has an on-going monthly cycle. Interviewers are required to make up to six calls at an address (an initial visit plus five 'call-backs'). In addition to the immediate reissue of contact sheets that have been wrongly completed or where the required number of call-backs has not been made, there is an on-going programme of reissuing 'non-contacts' in a bid to maximise the response rate. At the end of each fieldwork year a significant number of valid but 'non-contact' addresses remain 'live'.

The response rates for the SHS need to take account of the continuous nature of the survey. The data file for each year will contain a small proportion of interviews conducted on sample drawn the previous year. Similarly some of the addresses issued during any year will not be carried out until after the data file has been closed for analysis. These interviews are carried into the next data file. The response rates therefore report the outcomes for addresses sampled for a given period regardless of when the interview was carried out. Details of the most recent response rates are given at paragraph 4.2.1.

4 Fieldwork targets and outcomes

The requirements of the sample for the survey are expressed in terms of the ability of the sample to deliver the number of interviews required to meet particular requirements at particular points in time. There are also requirements in terms of the geographical scope of the interviews and their representativeness in terms of the types of people and areas included in the survey. Overall, the survey is expected to be an unbiased sample of Scottish households and adults and this is most likely to be achieved by maximising the level of participation in the survey.

With the sample designed to meet these objectives, these represent the key performance criteria for the survey. The survey's administration procedures are designed to minimise the impact of problems such as potential respondents not being at home or being unable to take part because of communication difficulties. Interviewers are required, for example, to make a minimum of 6 calls at each address on different days and at different times before it is considered 'no contact'. Even then, addresses will be reissued at a later stage in the fieldwork. Similarly 'soft' refusals such as 'too busy' or 'going out' are reissued.

Where interviewers are unable to conduct interviews in English or because the respondent is blind or partially sighted, these addresses can be revisited by an interviewer accompanied by an interpreter and, if appropriate, with showcards printed in a range of languages to minimise the extent to which language and communication barriers prevent people from taking part.

Nevertheless, participation in surveys is voluntary and some potential respondents refuse to take part. Similarly, no contact may be made at an address. This often reflects combinations of household types (single adults of working age), lifestyles (long working hours, active social lives), particular types of neighbourhoods and passive refusal (reluctance to open doors to strangers).

4.1 Sample performance

4.1.1 Ineligible addresses

The sampling for the survey makes assumptions about the proportion of addresses that will be ineligible for interview in each local authority. Ineligible addresses would include derelict, vacant and non-residential addresses. The allowance for ineligible addresses is based on the proportion that was actually identified in the course of the previous year's SHS fieldwork, the most recent two-year sweep when the sample was being designed in 2006. The extent to which these assumptions are accurate has an important bearing on the survey outcomes. If there are more 'deadwood' addresses, the interviewers have a smaller pool of addresses from which to achieve the target number of interviews. Conversely, a smaller proportion of 'deadwood' addresses should make it easier to achieve the target number of interviews but this target will

be met with a lower response rate. Thus, overall, if the proportion of deadwood differs from the sampling assumptions this might have some impact on achieving the interview target and the target response rate.

Table 4-1 shows the proportion of deadwood addresses assumed in each local authority when sampling for 2007-2010 and compares this with the proportion recorded by interviewers in the field from the 2009/2010 sample.

Table 4-1: Deadwood rate assumptions and actual deadwood
Sorted in descending order of deviation (absolute value) between actual and assumption

	Average deadwood 1999-2006	Actual deadwood 2009/2010	Deviation
	(%)	(%)	(% points)
Eilean Siar	8.1	18.1	10.0
Orkney Islands	6.5	13.4	6.9
South Lanarkshire	5.6	11.9	6.3
Perth & Kinross	9.5	13.4	3.9
Stirling	8.1	11.9	3.8
Shetland Islands	8.2	11.9	3.7
West Dunbartonshire	8.3	11.9	3.6
Moray	4.6	7.8	3.2
South Ayrshire	9.1	11.9	2.8
Aberdeenshire	5.8	8.4	2.6
West Lothian	9.5	11.9	2.4
Glasgow City	10.2	12.0	1.8
East Ayrshire	5.9	7.2	1.3
East Lothian	6.5	6.5	0.0
Highland	12.2	12.2	0.0
Scottish Borders	7.4	7.3	-0.1
East Dunbartonshire	4.2	4.1	-0.1
East Renfrewshire	4.4	4.0	-0.4
North Lanarkshire	7.9	7.2	-0.7
Edinburgh, City of	9.1	7.7	-1.4
Angus	9.9	8.1	-1.8
Fife	9.2	7.3	-1.9
Aberdeen City	9.8	7.3	-2.5
Renfrewshire	10.0	7.3	-2.7
Falkirk	9.3	6.3	-3.0
Dundee City	12.0	8.9	-3.1
Inverclyde	12.9	9.6	-3.3
Dumfries & Galloway	11.9	8.1	-3.8
Midlothian	10.0	5.1	-4.9
North Ayrshire	11.4	5.9	-5.5
Clackmannanshire	14.5	8.1	-6.4
Argyll & Bute	22.0	8.1	-13.9
All areas	**9.4**	**9**	**-0.4**

This shows that overall, and in many local authorities, the level of deadwood recorded by interviewers was close to that used as the basis for the survey sampling. There is, of course, some deviation from the assumptions, reflecting sampling variability in the data used for sampling and the sampled addresses. In spite of the deviation from assumptions, using different deadwood assumptions in individual local authorities rather than previous practice of a uniform 10% in all areas improves the structure of the sample and should contribute to meeting fieldwork targets. Where the experience differs from the assumptions this is likely to reflect a combination of factors such as:

- housing regeneration and redevelopment, which leads to demolition and vacant properties (increasing deadwood) and properties being brought back into use (lowering deadwood).
- expansion of holiday properties and second homes, which are ineligible for inclusion in the survey, raising deadwood.

4.2 Fieldwork performance

The profile of the sample and the level of deadwood are primarily qualities of the sampling frame and the assumptions used to sample. Inaccuracy and bias in these can have a knock-on effect on fieldwork performance. The other elements of fieldwork performance reflect:

- survey administration procedures and interviewer performance
- the availability of members of the public to be interviewed
- the ability of members of the public to participate in the interview
- the willingness of members of the public to participate in the survey.

Performance on each of these elements (as well as deadwood) is recorded as part of interviewers' attempts to secure interviews although there is, inevitably, interaction between these different aspects of performance.

4.2.1 Survey response rates

Overall, performance is summarised in the survey response rate and this is shown below for the sample drawn for 2009/2010.

Table 4-2: Summary of outcomes at issued addresses for 2009/2010 sample

	Frequency	Percent	Valid Percent
Complete interview	28,404	62.7	68.9
Office refusal	840	1.9	2.0
Refusal by selected respondent	5,953	13.1	14.4
No contact with anyone at the address	4,312	9.5	10.5
Broken appointment, no recontact	187	0.4	0.5
Refusal by proxy	413	0.9	1.0
Ill at home during survey period	427	0.9	1.0
Away/in hospital during survey period	270	0.6	0.7
Language	138	0.3	0.3
Other non-response	280	0.6	0.7
Total eligible for inclusion in the survey	**41,224**	**91.0**	**100.0**
Demolished/derelict	347	0.8	
Vacant/empty	1,988	4.4	
Not yet built/under construction	53	0.1	
Non-residential address	422	0.9	
Communal establishment/institution	86	0.2	
Address out of scope	633	1.4	
Insufficient address/no trace	347	0.8	
Other ineligible	205	0.5	
Total ineligible	**4,081**	**9.0**	
Total issued addresses	**45,305**	**100.0**	

In terms of assessing the performance of the survey fieldwork, contact was made and full interviews achieved at 69% of addresses. This represents an improvement over the outcome for the 2007/2008 survey period.

4.2.2 Trends in response rates

The response rate in 2009/2010 reflects a return to the long term response rate that has been achieved since the survey began in 1998. Response rates over the two-year period 2007/2008 were the lowest achieved on the survey since the 1999/2000 sweep and reflected the combined impact of the late start to 2007 fieldwork, errors in the scripts and the excessive length of the interview at the start of the 2007. Fieldwork performance improved in 2008 (response rates were 2% higher in 2008 than in 2007) and 2009 and 2010 have consolidated that improvement.

The table below shows the response rates for each local authority for each two-year period between 1999 and 2008 and compares the 2009/2010 rates with the average in each local authority in that period.

Table 4-3: Trends in SHS response rates 1999 to 2010

	1999/2000 (%)	2001/2002 (%)	2003/2004 (%)	2005/2006 (%)	2007/2008 (%)	Average 1999-2008 (%)	2009 /2010 (%)
Aberdeen City	65	67	66	66	64	65.6	63.3
Aberdeenshire	68	70	74	73	74	71.8	71.4
Angus	67	73	75	73	70	71.6	71.2
Argyll and Bute	71	69	73	74	76	72.6	69.3
Clackmannanshire	66	62	71	77	70	69.2	72.4
Dumfries and Galloway	69	69	73	72	72	71.0	72.7
Dundee City	62	66	67	69	64	65.6	69.7
East Ayrshire	69	71	75	71	69	71.0	69.6
East Dumbartonshire	68	69	73	69	71	70.0	68.9
East Lothian	67	63	67	67	65	65.8	68.7
East Renfrewshire	59	63	66	63	64	63.0	69.7
Edinburgh, City of	64	60	63	66	57	62.0	62.4
Eilean Siar	79	81	79	78	76	78.6	80.3
Falkirk	66	65	74	72	68	69.0	71.5
Fife	65	65	75	76	75	71.2	76.6
Glasgow City	62	63	60	59	57	60.2	60.4
Highland	68	71	70	71	71	70.2	70.3
Inverclyde	68	69	73	69	74	70.6	74.5
Midlothian	66	66	68	67	64	66.2	70.4
Moray	72	72	76	76	70	73.2	73.9
North Ayrshire	70	63	69	66	66	66.8	74.1
North Lanarkshire	61	64	67	70	64	65.2	66.5
Orkney Islands	70	80	80	77	83	78.0	81.8
Perth and Kinross	70	68	67	71	66	68.4	67.4
Renfrewshire	64	65	71	62	64	65.2	73.9
Scottish Borders	68	71	78	78	70	73.0	73.8
Shetland Islands	70	78	80	76	72	75.2	75.3
South Ayrshire	67	68	71	70	66	68.4	69.5
South Lanarkshire	64	65	67	67	64	65.4	69.2
Stirling	68	71	77	80	72	73.6	76.4
West Dumbartonshire	63	64	67	67	70	66.2	64.5
West Lothian	65	65	71	71	59	66.2	64.9
Total	**66**	**67**	**69**	**69**	**67**	**67.6**	**68.9**

Overall, between 1999 and 2010 the SHS response rate has been very resilient, fluctuating between 66% and 69% over the whole period. Even when comparing local authorities there is no clear trend, with wave-on-wave fluctuations but no obvious pattern of improvement or decline. The most notable differences between the 1999-2008 average and the 2009/2010 response rate are in East Renfrewshire, Renfrewshire, North Ayrshire and Fife where response rates in 2009/2010 are between 5 and 9 percentage points higher than the 1999-2008 average. The largest fall in response rates was in Argyll and Bute where the rate fell by 3 percentage points compared with the 1999-2008 average.

As has typically been the case since the survey began, the highest response rates are achieved in island and largely rural authorities. Response was highest in Orkney (84%), Eilean Siar (82%%) and Fife (78%).

The lowest response rates were, as they generally always have been, in Glasgow City, Edinburgh and West Lothian although each of these has increased compared with 2007/2008.

4.2.3 Achieved interviews compared with targets – household interviews

The number of interviews compared with the target, and the corresponding response rates, are the principal measures of survey performance although issues of data quality and bias also need to be considered. The table below compares interview targets and achievement in each local authority.

Table 4-4: Household interview targets and numbers achieved in each local authority, 2009/2010

	Target	Achieved	% of target achieved	Over / under achieved
Aberdeen City	1,060	1,057	100%	-3
Aberdeenshire	1,375	1,308	95%	-67
Angus	566	544	96%	-22
Argyll & Bute	693	678	98%	-15
Clackmannanshire	510	509	100%	-1
Dumf & Galloway	865	853	99%	-12
Dundee City	720	777	108%	57
East Ayrshire	657	608	93%	-49
East Dunbartonshire	510	472	93%	-38
East Lothian	611	639	105%	28
East Renfrewshire	500	535	107%	35
Edinburgh City	2,290	2,243	98%	-47
Eilean Siar	500	484	97%	-16
Falkirk	700	666	95%	-34
Fife	1,810	1,887	104%	77
Glasgow City	3,000	3,028	101%	28
Highland	1,282	1,233	96%	-49
Inverclyde	493	528	107%	35
Midlothian	556	565	102%	9
Moray	612	611	100%	-1
North Ayrshire	696	785	113%	89
North Lanarkshire	1,440	1,419	99%	-21
Orkney Islands	502	496	99%	-6
Perth & Kinross	795	784	99%	-11
Renfrewshire	820	900	110%	80
Scottish Borders	641	601	94%	-40
Shetland Islands	510	439	86%	-71
South Ayrshire	596	580	97%	-16
South Lanarkshire	1,390	1,422	102%	32
Stirling	563	569	101%	6
West Dumbartonshire	490	491	100%	1
West Lothian	756	693	92%	-63
	28,509	**28,404**	**100%**	**-105**

The highest levels of under-achievement were in Shetland Islands, Aberdeenshire, West Lothian, East Ayrshire and Highland. Taken together, these authorities account for 51% of the total shortfall (i.e. the sum of all the authorities where the interviews achieved is short of target). The highest percentage shortfalls were in West Lothian (92% achieved), Shetland Islands (86%) and East Ayrshire (92%).

4.2.4 Achieved interviews – the random adult

The two-part structure of the SHS interview requires the selection of a random adult within the household who completes the second half of the interview. This represents a second opportunity for potential

respondents to withdraw from the interview either because they refuse to take part or are unable, unavailable or not contactable for interview.

There is inevitably a degree of attrition between the household and random adult sections of the interview, especially where the person selected is not the same as the household respondent. This aspect of the survey has deteriorated since 1999/2000 when a random adult interview was achieved in 94% of households in which a household interview was completed. In 2009/2010, random adult interviews were achieved at 88% of households where a household interview was completed. This is the first time the random adult participation rate has dropped below 90%. The participation rate varied from 83% in East Renfrewshire to 97% in Orkney and 94% in Moray. An overall participation rate of 88% means that while the overall household response rate was 69%, the random adult response rate was 61%.

The combination of lower response rates to the household survey and attrition at the random adult stage means that the random adult response rates varied from 51% in Glasgow and 54% in Edinburgh and Aberdeen to 79% in Orkney and 73% in Eilean Siar.

Table 4-5: Random adult (RA) response rates, 2009/2010

	Valid addresses	Household interviews	Random adult interviews	RA interviews as % of valid addresses	RA interviews as % of household interviews
Aberdeen City	1,669	1,057	905	54%	86%
Aberdeenshire	1,831	1,308	1,190	65%	91%
Angus	764	544	495	65%	91%
Argyll and Bute	978	678	623	64%	92%
Clackmannanshire	703	509	434	62%	85%
Dumfries and Galloway	1,173	853	764	65%	90%
Dundee City	1,115	777	658	59%	85%
East Ayrshire	874	608	525	60%	86%
East Dumbartonshire	685	472	422	62%	89%
East Lothian	930	639	573	62%	90%
East Renfrewshire	768	535	445	58%	83%
Edinburgh City	3,593	2,243	1,950	54%	87%
Eilean Siar	603	484	441	73%	91%
Falkirk	931	666	561	60%	84%
Fife	2,465	1,888	1,667	68%	88%
Glasgow City	5,013	3,028	2,540	51%	84%
Highland	1,753	1,233	1,091	62%	88%
Inverclyde	709	528	467	66%	88%
Midlothian	802	565	483	60%	85%
Moray	827	611	574	69%	94%
North Ayrshire	1,060	785	722	68%	92%
North Lanarkshire	2,134	1,419	1,256	59%	89%
Orkney	606	496	479	79%	97%
Perth and Kinross	1,164	784	670	58%	85%
Renfrewshire	1,218	900	804	66%	89%
Scottish Borders	814	601	539	66%	90%
Shetland	583	439	406	70%	92%
South Ayrshire	834	580	538	65%	93%
South Lanarkshire	2,055	1,423	1,223	60%	86%
Stirling	745	569	521	70%	92%
West Dumbartonshire	761	491	427	56%	87%
West Lothian	1,064	691	589	55%	85%
	41,224	**28,404**	**24,982**	**61%**	**88%**

5 Weighting

Two types of weighting are potentially necessary with a sample of this kind. The first is intrinsic to the survey design and represents weights necessary to compensate for unequal probabilities of selection for individuals, households or other units of analysis. The second may be necessary to counteract the effects of non-response bias. Although these represent two distinct rationales for weighting, in terms of analysis the different weights are combined into a single weighting variable for each unit of analysis.

In addition to these weights, grossing can also be calculated which allow the survey results to be expressed in terms of estimates of the number of people from the population rather than a percentage.

In addition to these, there are weights associated with the sub-sample that comes from the selection of a random schoolchild.

Finally, the SHS recently introduced additional weighting procedures to address known areas of bias in the survey sample. These calibration procedures have been applied to the main sample and are discussed further below, but they result in two further weights for both the household questions and the random adult questions – one which calculates weighted survey estimates and one which calculates grossed up population estimates.

In the SHS, there are 10 weights that can be used but two of these – LA_WT and IND_WT – are used for most analyses, with the others used for smaller specific subsets of the sample and the 'old' basic design weights retained for analysis that omits the population calibration. The table below shows the different types of weights available for the survey.

	Basic 'design' weights	Calibrated weights	Grossing to population estimates
Main sample			
Household	LA_OLDWT	LA_WT	LA_GRWT
Random adult	IND_OLDWT	IND_WT	IND_GRWT
Random schoolchild	KID_OLDWT	KID_WT	KID_GRWT
Travel diary	-	TRAV_WT	-

Although LA_WT and IND_WT use the names of the main weight variables calculated for previous waves of the SHS, it should be noted that these variables are now the outcome of recently introduced calibration processes. They have been given these names to help ensure that they become the default for analysis of the SHS data. The previous design weights remain in the data and have been renamed LA_OLDWT and IND_OLDWT.[8]

[8] Calibrated weights have also been produced for each of the preceding two-year sweeps. Revised weights can be obtained from the SHS website.

- LA_WT is used for analysis of data about the household and data collected from or about the HIH and spouse in the main SHS sample. This includes all variables asked in the first part of the interview, apart from the questions about the random schoolchild and the random child receiving childcare. This weight is associated with two others: LA_OLDWT which is the previous design weight and LA_GRWT which is the calibrated grossing weight required to provide whole population estimates.
- IND_WT which is used for analysis of data in derived variables about the random adult or collected from the random adult. This includes all variables in the second part of the interview. This weight is associated with IND_OLDWT and IND_GRWT.
- KID_WT which is used for analysis of questions related to the random schoolchild – HE9 to HE17N inclusive (see ***Questionnaire***). Like LA_WT, this is now calibrated to population totals and is associated with KID_OLDWT, which is the previous design weight and KID_GRWT which is the calibrated grossing weight to provide population estimates.
- TRAV_WT, contained in the travel diary data, which is used for analysing that data.

5.1 Weighting for analysis based on household data

The weight for analysis of household data, LA_WT, has three main elements. Firstly, it is necessary to 'weight up' those local authorities which were under-sampled and 'weight down' those which were over-sampled (this is a weight of the first type mentioned above, which adjusts for unequal probabilities of selection). Secondly, the weight addresses any additional disproportionality introduced by response rates differing from the target for each local authority.[9]

Weights are calculated for each local authority so that in each quarterly data file, the data is nationally representative. This should allow any published findings to be reproduced by selecting the relevant quarter's data. In practice, however, it may not be possible to reproduce exactly some of the results from earlier publications if the data for that quarter were subsequently changed (e.g. to correct errors that were identified later) and because there is some overlap between the quarter in which interviews take place and the quarter's data with which it is processed. For example, the data processed as Q4 2009 contained data from interviews carried out in the first quarter of 2010 so although they were weighted as Q4, they have a value of 1 for the Quarter variable.

[9] Strictly speaking, the first component of this is only approximately a design weight. Since deadwood rates vary by local authority the differential sampling would need to compare the population distribution with the distribution of valid addresses. The second part crudely deals with differential response rates in the sense that it assumes that actual deadwood = assumed deadwood and that therefore any variation in the achieved interviews from the target number of interviews is a reflection of differences in responses rates from those used in the sampling. However, we have already seen that this deadwood assumption although broadly realistic is not always true. The reason for weighting in this way is practical. At the end of each quarter, the status of each sampled address is not known because non-contacts and refusals might be reissued and might become either deadwood or interviews so we can do neither the design element (because we don't know the number of valid address) nor the non-response element (because we don't know the final status of each address). However, by looking at the distribution of achieved interviews across local authorities and weighting so that each LA represents the correct proportion of the weighted sample, we can adequately address both to meet the requirement of a nationally representative sample each quarter.

No other weight is applied across all cases in order to adjust for the unequal probabilities of selection. Strictly speaking, however, a corrective weight should be applied in those cases in which the Multiple Occupancy Indicator (MOI) on the Postcode Address File (PAF) is found to be inaccurate. The reason for this is that a property-type bias might otherwise be introduced. For example, if tenement properties were consistently found to contain multiple dwellings when the MOI had indicated that they contained just one, each achieved interview at such an address should be given a weight proportional to the actual number of dwellings, to compensate for the reduced probability of selection for each dwelling at that address. All properties within that local authority area should then be weighted back down slightly in order that the actual and weighted sample sizes remain the same.

In practice, the MOI has been found to be inaccurate in only about 2% of cases. The impact of weighting to correct for these would have been negligible so it was decided not to weight by the MOI in order to avoid additional complexity in the weighting scheme for the survey.

Similarly, in theory an additional weight should be applied in cases where a dwelling contains more than one household, only one of which is interviewed, in order to adjust for the lower probability of selection for each of the households in that dwelling. In practice, however, as only a very small number of dwellings were found to contain more than one household, the use of such a weight would make very little difference to the overall results, and it was therefore felt that it was not worthwhile introducing further complication to the weighting calculations.

The calibration weights and grossing weights were designed following a review of the weighting strategy for the survey. Since the start of the survey, the weighting scheme for the SHS had been kept intentionally simple. This reflected, in part, a desire to keep the processes of the survey straightforward so that the data can be made available for analysis as quickly as possible. It also reflected the limited extent to which the SHS data differs substantially from comparator data, as we discuss below.

This aspect of the survey has been subject to review by the Office for National Statistics as part of a major study comparing non-respondents to the SHS with Census data. This study concluded that while comparison with the Census showed some bias in the SHS, this was not substantial although some corrective weighting was recommended. Further work looking at the scope for corrective weighting was undertaken in 2008 and 2009 and developed further with revised weighting arrangements developed for the 2007-2010 phase of the survey. Further details of the weighting review, the conclusions reached and the impact of the revised weighting can be found at www.scotland.gov.uk/SHSMethodology.

The revised weighting uses established calibration techniques to fit the composition of responding households to mid-year population estimates produced by the General Register Office for Scotland. The weighting is carried out using g-Calib – a module for SPSS produced by Statistics Belgium.[10]

[10] Information about calibration weighting and g-Calib can be found at http://www.statbel.fgov.be/studies/paper03_en.asp

5.2 Weighting for analysis based on individual (random adult) data

Using the Postcode Address File produces a sample of households, so for analysis of individual level data it is also necessary to weight the responses of the random adult by the number of adults resident in the household who were eligible for interview.[11] The reason for this is that individuals living in larger households have a lower probability of selection than adults in, for example, single adult households where that one person must be sampled.

As a result of this, the *unweighted* profile of 'random adult' respondents will tend to be skewed towards those sections of the population most likely to live in households with fewer adults (older people and older females in particular) and away from those likely to live in households with larger numbers of adults (younger people). Once the data are weighted by the number of eligible adults in the household, however, one should see the profile correct itself significantly. In most surveys of this kind, however, some under-representation of younger people and males, and over-representation of older people and females, is likely to remain because of the effects of non-response bias.

Analysis of data based on the random adult also requires a further weight to take account of differences between the number of such interviews completed in each local authority area and the actual adult population of such areas. Like the element of the household data weight which adjusts for differences in fieldwork outcomes by local authority, this is intended not to compensate for unequal probabilities of selection but to ensure that the final profile of 'individual' data correctly reflects the relative populations of the different local authority areas once variations in fieldwork outcomes have been assessed. This is not identical to the weight described for analysis of household data, since variation in response rates for the second part of the interview may have produced a slightly different distribution from that of 'householder' interviews.

Like the household weighting, this basic 'design' weighting is then extended by comparing the age/sex profile of the adult respondents with the corresponding profile of the population within each local authority. The intention is to address the under-representation of some population segments, especially younger people in general and younger men in particular. The resulting calibrated weight (IND_WT) should be used for analysis of all random adult data.

5.3 Weighting for analysis based on the 'random schoolchild'

Data relating to the information collected about a 'random schoolchild' needs to be weighted so that this information will represent correctly the population of schoolchildren resident within households. If not, it will proportionately over-represent the characteristics and experiences of 'only' children and under-represent those of children from larger families. The weight for the random schoolchild case is created by combining the number of schoolchildren in the household and the relevant local authority weight. The weighted

age/sex profile of pupils is compared with the profile of all pupils, within local authorities and the weights calibrated to ensure that the survey data matches the population profile.

The major difference in the calibration of the schoolchild data is that while LA_WT and IND_WT are based on official population estimates, the population of schoolchildren is created within the survey by looking at all household members identified as schoolchildren and using the household grossing weight to estimate the number of pupils in each age/sex group within each local authority. At a national level, the grossed survey estimates are very close to the most recent official estimate. The 2007/2008 survey estimates that there were 724,116 schoolchildren in Scotland. The 2007 schools census estimates that there were 723,196 schoolchildren across both the local authority and independent sectors.

5.4 Weighting for analysis based on the Travel Diary

Examination of the SHS data suggests that significantly fewer interviews take place on Fridays, Saturdays and Sundays than on other days of the week. As differences in the proportions of adults interviewed on each day of the week will affect the Travel Diary data's representativeness of travel patterns for the week as a whole, it was decided to introduce a weight to compensate for this. This simply 'up-weights' interviews carried out on days of the week on which fewer than one-seventh of all interviews have taken place and 'down-weights' those carried out on days on which more than one-seventh of all interviews have been completed.

It is also apparent that the distribution of interviews by the day of the week differs for certain sub-sections of the adult population. For example, disproportionately more adults in full-time employment are interviewed at the weekend (due to their greater availability then), thus yielding an inaccurate picture of the travel patterns of those in full-time employment. The Travel Diary weighting factor is therefore refined to compensate for this.

The weight created for any analysis of the Travel Diary combines the above weighting factors and the calibrated random adult weight. Further information about the Travel Diary, including a comparison to the National Travel Survey, is available in the Travel Diary Background Information report.[12]

[11] This weight incorporates the local authority weight described earlier. This is necessary for all analyses (whether of households or individuals) if the Scottish population resident in private households is to be represented accurately. The way in which weights are combined is further described later in this section.

[12] http://www.scotland.gov.uk/Topics/Statistics/Browse/Transport-Travel/TDback

6 Limitations of the data

There are a number of important methodological and data issues that users need to be aware of when using the SHS data.

Like all sample surveys, the SHS can only produce estimates and these estimates are limited by a number of factors.

- Sample coverage – although there are no geographical exclusions to the survey, the sampling frame does not cover the whole population because of a combination of inherent limitations and administrative errors and delays.
- Sampling variability – all samples can differ from the population by chance. This is often referred to as sampling error.
- The number of cases that analysis is based on – estimates based on large samples are more accurate than those based on small samples.
- Bias in the achieved sample – if a sample under-represents sections of the population or if a large proportion of people do not answer some questions, the estimates may differ substantially from the population for reasons that are not a result of chance. For example, in 2009, the unweighted sample of adults is 56% female and even after basic weighting 55% of the sample is female, but the true figure in the population is only 52%.[13] This is an example of bias caused by young males, in particular, being difficult to contact or refusing to take part in the survey.

Although the introduction of calibration weighting addresses the disparity between the age/sex composition of the sample and the known composition of the population, it does so on the assumption that respondents and respondents do not differ in terms of survey measures that do not form part of the weighting. The review of the weighting strategy generally found that calibration brought the survey estimates closer to census estimates but like all surveys, the potential for bias remains a limitation that should be considered.

The SHS is also limited in the amount of detail it can collect about some topics. For example, it was not designed to provide reliable "economic" statistics (e.g. employment/unemployment rates and average earnings).

The SHS's information about the **economic status** of members of the household reflects the view of the respondent to the "household" part of the interview, and so may not conform to official definitions of

[13] General Register Office for Scotland mid-year population estimates June 2010

employment and unemployment, for example. As a result, the SHS cannot provide estimates of unemployment that are comparable to official statistics of unemployment.[14]

There are several reasons why the SHS data on **income** may not be completely accurate.

- The SHS only collects information from, or about, the Highest Income Householder and, if there is one, their spouse or partner.
- Information is provided "off the top of the head" as part of an interview on many other topics. There is no requirement to refer to pay slips or bank statements to check the figures.
- Some people may not know the correct figure (particularly in the case of the income of a spouse/partner), and may just provide a guess, perhaps based on a level that they remember from some time ago.
- Other interviewees may under-state their income because they do not want to reveal how much they really earn.
- Because about a third of the households in the sample are unwilling or unable to provide income information, values for some or all of the main components of income have to be imputed.[15]

In 2004, researchers commissioned by the Scottish Executive and Communities Scotland compared the income data collected by the SHS and the Scottish House Condition Survey (SHCS) with the income statistics produced from the Family Resources Survey.[16] Their main conclusions were:

- the SHS (and SHCS) under-estimate total household income, due to collecting only the income of the highest income householder and any spouse/partner
- when households with one adult or two adults who are spouses/partners are compared, there is good agreement between the SHS/SHCS and FRS income distributions for such households
- SHS (and SHCS) greatly under-estimate investment income and interest payments compared to FRS
- uncorrected bias in the SHS (and SHCS) age and sex distributions affects income distributions, particularly for one person households
- overall income from benefits agrees well between the surveys, but the individual benefits may be less accurately classified in the SHS (and SHCS).

As a multi-purpose survey of households, the SHS is not designed to provide the kinds of information about economic activity and household income that can be obtained from more specialised surveys such as the Labour Force Survey and the Family Resources Survey, which have questions and procedures

14 See, for example, the comparison of economic activity estimates from the SHS and the Annual Scottish Labour Force Survey in section 7 'Bias and Data quality'.

15 Details of the imputation process can be found in the SHS Annual Report *Scotland's People: results from the 2007/2008 Scottish Household Survey*

16 Raab, G., MacDonald, C. & Macintyre, C. (2004) *Comparison of income data between surveys of Scottish households*. Research commissioned by Communities Scotland. Electronic copies are available on the SHS website.

which are designed to obtain much more reliable information on those matters than the SHS can collect. The SHS has questions on such topics *only* for selecting the data for particular groups of people (such as the unemployed or the low-paid) for further analysis, or for use as "background" variables when analysing other topics (such as the means of travel or the frequency of driving).

Although the SHS has a large sample that covers the whole of Scotland, it has some geographical limitations because of the sample sizes in small local authorities and because it is designed to be representative only at national and local authority level. This means:

- users need to be mindful of the sampling errors for analysis but especially when this is based on breakdowns within a single local authority
- it is not appropriate to undertake geographical analysis below local authority level since the sampling techniques used in some local authorities cannot guarantee representativeness in smaller areas.

6.1 Quarterly data for Scotland as a whole

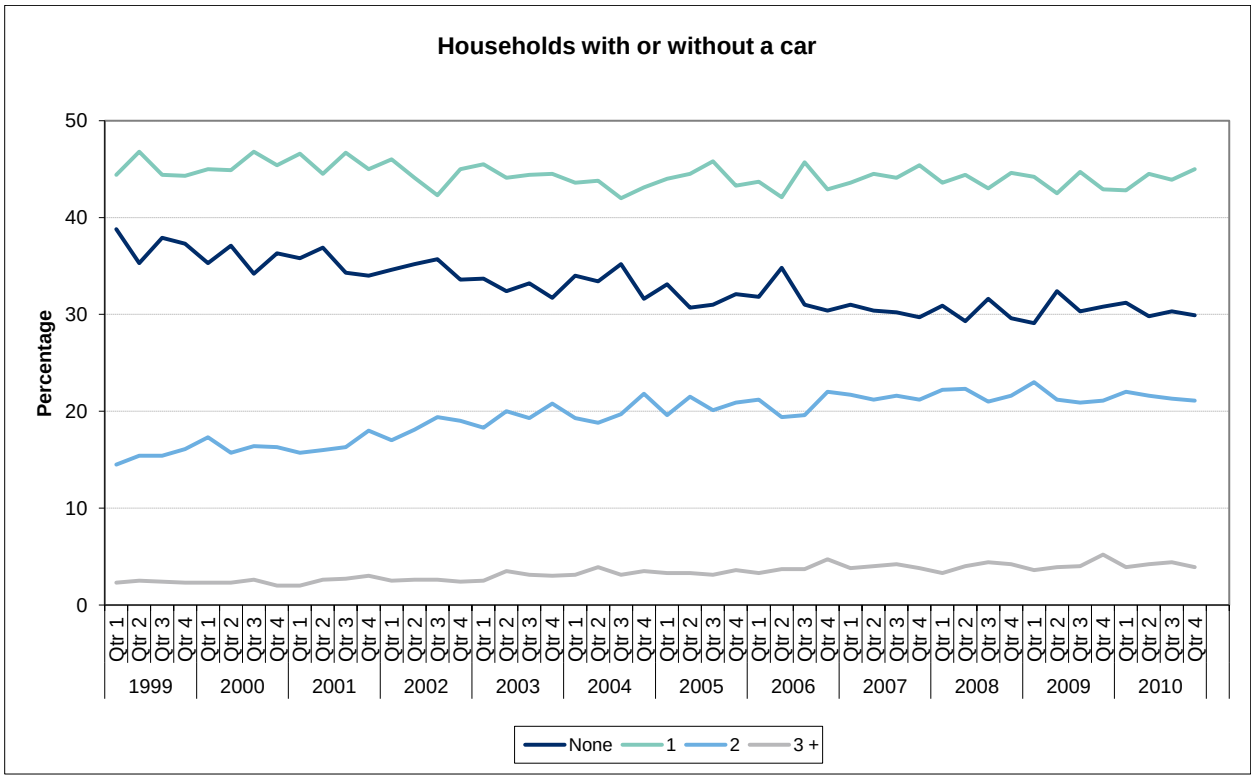
The SHS was designed to provide results which are representative for Scotland as a whole for each quarter of the year. Although based on a large sample (nearly 4,000 households per quarter), they are still subject to sampling errors, so may well fluctuate from one quarter to the next. Therefore, apparent quarter-to-quarter changes should be interpreted cautiously, as they may well be due to sampling variability rather than representing genuine change.

This can be seen if one looks at the apparent quarter-to-quarter changes in some figures which one would expect to change only gradually from one quarter to the next - especially figures which show trends that one would *not* expect to be subject to short-term reversals. The SHS's Quarterly Data Release[17] provide a set of quarterly tables and charts. Examples of two of these appear on the following pages. The first example shows, quarter-by-quarter since the survey started, the (weighted) percentages of households in the sample with various numbers of cars available for private use; the second shows quarter-by-quarter figures for household tenure.

In both cases, the quarterly charts show the kinds of long-term trends that one would expect (e.g. a gradual increase in two-car households) - but with some apparent "wobbliness" in the lines. Given the nature of car ownership and household tenure, one would not expect sudden short-term departures from the long-term trend (such as a sharp fall in the percentage of homes which are owned outright). However, the survey results sometimes suggest very surprising quarter-to-quarter changes. For example, the first chart shows that the (weighted) percentage of households with 3+ cars appears to vary from quarter to quarter. The cause cannot be any real change in car ownership across Scotland: it must just be sampling variability (the "luck of the draw" regarding which households were included in the sample in each quarter, and which of them agreed to take part in the survey). In the Annual Report's Appendix on confidence intervals and statistical significance, Table A4.1 indicates that the 95% confidence limits for an estimate of

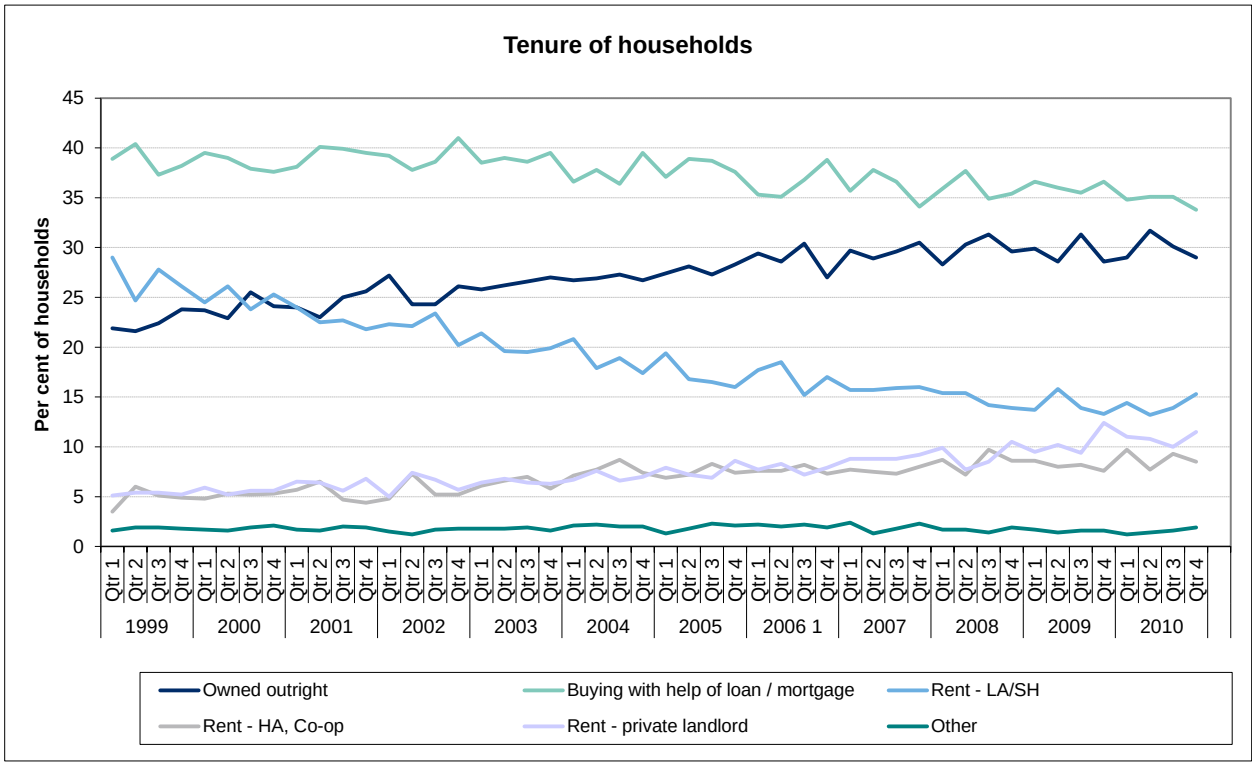
5% based on a sample of 4,000 cases are about +/- 0.8%. The apparent fluctuation in these percentages is a good illustration of such sampling variability.

Examples of charts and tables showing quarterly figures



Apparent quarter-to-quarter changes in the figures must be interpreted with caution: they could well be the result of sampling variability.

[17] Available at http://www.scotland.gov.uk/Topics/Statistics/16002/DataTrends



Apparent quarter-to-quarter changes in the figures must be interpreted with caution: they could well be the result of sampling variability.

7 Bias and data quality

The issue of bias arises in every survey of the population. There are a number of sources of bias, some of which reflect aspects of the survey design (such as the sampling frame or who is deemed eligible for interview). However, bias is also a reflection of those aspects of fieldwork outcomes mentioned above:

- the quality of survey administration procedures
- whether potential respondents can be found at home at times when interviewers call
- whether they are able to participate i.e. not restricted by ill health, disability or communication barriers
- the willingness of members of the public to participate in the survey.

A high response rate is generally viewed as one of the key measures of data quality and, all other things being equal, a high response rate and a large sample should ensure accurate estimates. However, to the extent that non-response to the survey is not spread evenly, either geographically or between sub-groups of the population, the resulting bias will limit the accuracy of the survey's estimates. The revisions to the survey's weighting strategy (see section 5 Weighting) are intended to minimise the extent of bias and in addition to that the question of residual bias is considered by comparing key results from the SHS with comparator data. Since the publication of the 2001 Census, this source is the most accurate comparator for population data and in spite of being a few years behind the current SHS, population measures such as age distribution and household types change little from year-to-year.

Since the expectation is that most users of the SHS will be interested in the main survey rather than the Culture and Sport module, consideration of data quality is restricted to the main survey.

7.1.1 Household type, property type, tenure and number of bedrooms

Compared with the 2001 Census, single adult and large adult households are now slightly over-represented, and single pensioner and older smaller households under-represented, reversing what has been observed in the sample since 1999. This is partly a reflection of the revised weighting and improvements in response rate but perhaps also reflects real underlying change, given that the Census is now eight years old. (Table 7-1).

Table 7-1: Comparison of household types in the 2001 Census and the 2009/2010 SHS (%)

	2001 Census	2009/2010 SHS *
	(n=2,192,246)	(n=28,404)
Single adult	17.9	18.7
Small adult	16.9	19.0
Single parent	5.6	5.0
Small family	13.3	12.7
Large family	7.1	6.4
Large adult	11.2	9.9
Older smaller	13.0	14.1
Single pensioner	15.0	14.2

* SHS data weighted by LA_WT

As Table 7-2 shows, the sample appears robust in terms of the variables associated with accommodation/property characteristics. Compared with the 2001 (which is eight years older than the data in the 2009 SHS) there is a higher proportion of houses and a lower proportion of flats. Outright ownership appears to be over-represented compared with owning with a mortgage and in total, owner-occupation is over-represented in the SHS compared with rented tenures.

These features are likely to reflect long-term trends in housebuilding, with a preference for houses rather than flats, and changes within areas subject to urban renewal where many high rise flats and public sector flats have been demolished and replaced by terraced housing. Similarly, there has in recent years been a significant move towards local authority stock being transferred to housing associations and other public landlords, and this is often combined with urban renewal, resulting in changes in both the form and tenure of dwellings.

Table 7-2: Comparison of housing variables in the 2001 Census and the 2009/2010 SHS

	2001 Census (n= 2,192,246)	**SHS** (n=28,404)
	%	%
Property type* ‡		
House or bungalow	64	67
Flat, Maisonette or Apartment	35	33
Other	1	1
Tenure* †		
Own outright	23	30
Own with mortgage	39	36
Rent	35	33
Local authority	*22*	*14*
Housing Association/Co-operative	*6*	*8*
Private rented	*7*	*11*
Other	4	2

* SHS data weighted by LA_WT
‡ includes households in shared dwellings
† Pays part rent and mortgage (shared ownership) included in 'Own with mortgage'

7.1.2 Age and sex profile of the 'random adult' sample

When a single adult is randomly selected within households, the unweighted sample of adults always under-represents those living in multi-adult households, since each has a smaller chance of selection for interview. As Table 7-3 shows, weighting to equalise probabilities of selection generally has the effect of bringing the profile of the 'random adult' sample closer to that of the adult population. The SHS data shown have been weighted both by the number of adults resident in the household and by the local authority weight described in the previous section. These two weights tend to act in the same direction, since those larger local authority areas which are 'weighted up' also tend to be ones with a higher average household size.

Table 7-3: Comparison of weighted and unweighted age and sex profile of 2009/2010 SHS data with 2010 GROS mid-year estimates

	Population estimates for 2010	SHS random adults Unweighted	SHS random adults weighted*	SHS random adults weighted**
	%	%	%	%
Male				
16 – 24	7.4	3.7	5.2	7.4
25 – 59	28.1	24.5	25.6	28.0
60 plus	12.4	15.5	14.6	12.4
Total	**47.9**	**43.7**	**45.5**	**47.8**
Female				
16 – 24	7.1	4.6	5.8	7.1
25 – 59	29.4	30.0	30.9	29.4
60 plus	15.6	21.6	17.8	15.7
Total	**52.1**	**56.3**	**54.5**	**52.2**
All adults		(n=24,982)	(n=24,982)	(n=24,982)
16 – 24	14.5	8.2	11.1	14.5
25 – 59	57.5	54.6	56.5	57.3
60 plus	28.0	37.2	32.4	28.1
Total	**100.0**	**100.0**	**100.0**	**100.0**

* Weighted by number of adults and local authority size (IND_OLDWT)

** Weighted by IND_WT

However, even after design weighting has been applied (the third column of figures), the weighted random adult sample for 2009 still does not match the profile of the adult population suggested by the GROS estimates with, as expected, under-representation of younger people in general and 16-24 year olds in particular. Consequently, older people are over-represented in the random adult sample.

The calibration weighting brings the age/sex profile of the sample very close to the profile of the adult population, with the very small differences a reflection of the impact of rounding at a local authority level and summation across local authorities.

7.1.3 Ethnicity

The ethnicity questions were changed in June 2007 so that rather than collecting information about all household members, only the ethnicity of the random adult was collected. This means that although Table 7-4 compares the SHS profile with that of the Census, the comparison is not comparing like with like: the Census shows the profile of the whole population while the SHS shows the estimated profile of the adult population. Nevertheless, the figures suggest that there is a reasonably good agreement between the Census and the SHS and that the SHS appears to picking up on recent phenomena like the migration to Scotland of adults from East European countries. This is reflected in the higher proportion in the 'Other white' category.

Table 7-4: Comparison of ethnicity in Census 2001 and 2009/2010 SHS

	% of Census population 2001	% of adults 2009/2010 SHS
White	**98.0**	**96.7**
Scottish	88.1	81.5
Other British	7.4	11.7
Irish	1.0	0.7
Any other White background	1.5	2.8
Mixed	**0.2**	**0.2**
Any mixed background	0.2	0.2
Asian, Asian Scottish or Asian British	**1.3**	**2.0**
Indian	0.3	0.5
Pakistani	0.6	0.7
Bangladeshi	0.0	0.1
Chinese	0.1	0.4
Any other Asian background	0.3	0.4
Black, Black Scottish or Black British	**0.1**	**0.5**
Caribbean	0.0	0.0
African	0.1	0.3
Any other Black background.	0.0	0.1
Other ethnic group	0.2	0.4
Refused / don't know		0.1

7.1.4 Urban Rural classification

Analysis of the Scottish Household Survey makes extensive use of the Scottish Government's Urban Rural classification of assigning areas into different degrees of urbanity and rurality. This classifies settlements according to their size and for settlements with a population of less than 10,000, their proximity to a settlement with a population of 10,000 or more.[18]

Table 7-5 compares the urban/rural classification of the SHS sample for 2009/2010 with the profile of all addresses sampled for the survey, the profile of eligible addresses and participating households. This shows that the addresses sampled in 2009/2010 (column 2) under-represent urban areas and over-represent rural areas (as the sample is designed to do) but when disproportionate sampling is taken into account by weighting, the profile matches the population.

[18] http://www.scotland.gov.uk/Topics/Statistics/About/Methodology/UrbanRuralClassification

Table 7-5: Comparison of all Scottish households, all sampled households, all eligible households and participating households by urban rural classification

	All Scottish addresses*	All sampled addresses (unweighted)	All sampled addresses**	All eligible households**	All participating households***
Large urban areas	40	38	40	40	40
Other urban	30	28	30	30	30
Small accessible towns	9	8	9	9	9
Small remote towns	4	5	4	4	4
Accessible rural	11	11	11	11	11
Remote rural	6	9	6	6	6

* Weighted by number households within each unit postcode

** Weighted to reflect disproportionate sampling across local authorities

*** Weighted to reflect disproportionate sampling and non-response across local authorities

Comparison of eligible addresses and sampled addresses suggests that ineligible addresses are fairly evenly spread across the urban rural categories and that the variations in assumptions about the level of ineligibility helps to retain this balance. The households at which SHS interviews were achieved and the classification of all households sampled at a local authority level shows that there is a good match between the two within local authorities.

7.1.5 Economic activity

One area where the results of the SHS indicate significant differences from other sources is in relation to indicators of economic activity. As the following table shows, the most recent results from the Labour Force Survey (LFS) suggest that the SHS may be under-representing people in employment, and over-representing the economically inactive. It should be emphasised, however, that the information from the SHS shown here is based on the respondent's own classification of their economic activity (collected at the start of the interview)[19], rather than on the full International Labour Organisation definition, which is not classified by the respondent and is the basis for official estimates of unemployment. So the differences might simply reflect differences in interpretation. For example, students with part-time jobs might report their main activity as 'in higher or further education' in the SHS and be classified as economically inactive, whereas the LFS would categorise them as economically active and count them as employed. The SHS is not an official source of statistics on employment (see section 6 on limitations of the data).

[19] This is further complicated by the fact that where the household respondent is not subsequently selected as the random adult, the classification is given by the household respondent and attributed to the random adult.

Table 7-6: Comparison of economic activity variables among adults of working age

	2008 Annual Population Survey	2009/2010 SHS
	%	
Males	(n=14,241)	(n=8,114)
Employed	78.7	70.2
Unemployed	4.4	9.1
Economically inactive	16.9	20.8
Females	(n=14,294)	(n=9,833)
Employed	72.3	62.2
Unemployed	3.6	4.4
Economically inactive	24.1	33.4
All adults	(n=28,535)	(n=17,947)
Employed	75.6	66.1
Unemployed	4.0	6.7
Economically inactive	20.4	27.2

* weighted by number of adults and local authority size

Figures in this table have been calculated using all working age people (16-64) as the denominator, headline unemployment statistics are not calculated on this basis

Annual Population Survey data are sourced from quarterly Labour Force Survey data and the annual Labour Force Survey boost data.

8 Survey design factors and complex standard errors

Data collected in surveys are always an estimate of the true proportions in the population. The accuracy of these estimates – the sampling error – can be calculated for any estimate in the survey using information about the proportion of people giving the response and the number of people in the sample (or sub-sample). The sampling error can be expressed as a 'confidence interval', which can be added to and subtracted from the survey estimate to give a range within which it is fairly certain that the true value lies.

Since the SHS is not a simple random sample (SRS) design, the confidence intervals need to take account of the impact of clustering and stratification. The SHS, therefore, has what is known as a 'complex standard error'. While for some variables the design of the sample improves the precision of the survey estimates compared with a simple random sample, the overall effect of the survey design is to reduce the precision of the estimates. The relationship between the complex standard error and the theoretical simple random sample standard error for a sample of the same size is summarised in the 'design factor'. The expresses the extent to which the design of the survey broadens (or occasionally narrows) the confidence intervals for particular estimates.

The Taylor Expansion Method was used to calculate the complex standard errors for a series of results in the study. This is a well-established technique for working through the effects of stratification and clustering. As can be seen from Table 8-1, these ranged from 0.86 to 1.31. The overall average is 1.18, but that should not be taken as a 'typical' value, given the distribution of values across different variables. However, it suggests that the original assumption of a design effect of 1.1-1.2 was not unreasonable and using a value of 1.2 as a 'rule of thumb' for adjusting the standard errors of the survey data would account for the design factors associated with most variables in the survey.

The 95% confidence intervals shown are based on complex standard errors.

Table 8-1: Design factors and confidence intervals for key variables in 2009/2010 data

Characteristics	Estimate	95% Confidence Intervals		SRS error for the same size of sample	SHS Complex Standard Error	Design Factor
Tenure						
Owner-occupied	65.4	64.8	66.1	0.28	0.33	1.16
Social-rented Sector	22.5	21.9	23.0	0.25	0.30	1.20
Privately rented	10.6	10.2	11.0	0.18	0.21	1.16
Property type						
House	66.6	65.9	67.3	0.28	0.36	1.29
Flat/maisonette	32.8	32.1	33.5	0.28	0.36	1.29
Economic status of working age adults						
Full time employee	36.2	35.4	36.9	0.30	0.37	1.23
Part time employee	11.0	10.6	11.5	0.20	0.24	1.22
Self-employed	5.8	5.4	6.2	0.15	0.19	1.28
Unemployed	5.4	5.0	5.7	0.14	0.19	1.31
HIH or partner has a bank/ building society account	92.6	92.2	93.0	0.18	0.20	1.10
Marital status of all adults						
Married/cohabiting	51.1	50.3	51.8	0.32	0.39	1.23
Separated/divorced	9.5	9.1	9.8	0.19	0.18	0.99
Single/never married	31.9	31.2	32.7	0.29	0.38	1.29
Widowed	7.5	7.3	7.8	0.17	0.14	0.86
Access to the internet	66.8	66.1	67.5	0.32	0.35	1.08
Travel to work in a car	59.9	58.9	60.9	0.45	0.52	1.16
Reporting long-standing illness, disability or health problem	23.8	23.2	24.4	0.27	0.31	1.16
Rating of area as a place to live						
Very/fairly good	93.5	93.2	93.9	0.16	0.19	1.22
Not good	6.5	6.1	6.8	0.16	0.19	1.22
Smoking Status of all adults						
Smokes	24.2	23.6	24.9	0.27	0.33	1.20
Does not smoke	75.8	75.1	76.4	0.27	0.33	1.20

HIH = Highest income householder

A NATIONAL STATISTICS PUBLICATION FOR SCOTLAND

Official and National Statistics are produced to high professional standards set out in the Code of Practice for Official Statistics at http://www.statisticsauthority.gov.uk/assessment/code-of-practice/code-of-practice-for-official-statistics.pdf. Both undergo regular quality assurance reviews to ensure that they meet customer needs and are produced free from any political interference.

Statistics assessed, or subject to assessment, by the UK Statistics Authority carry the National Statistics label, a stamp of assurance that the statistics have been produced and explained to high standards and that they serve the public good.

Further information about Official and National Statistics can be found on the UK Statistics Authority website at www.statisticsauthority.gov.uk

SCOTTISH GOVERNMENT STATISTICIAN GROUP

Our Aim
To provide relevant and reliable information, analysis and advice that meet the needs of government, business and the people of Scotland.

For more information on the Statistician Group, please see the Scottish Government website at www.scotland.gov.uk/statistics

Correspondence and enquiries

Enquiries on this publication should be addressed to:

SHS Project Team
Communities Analytical Services
1-F (Dockside) Victoria Quay
Edinburgh EH6 6QQ
Telephone: 0131 244 8420;
Fax: 0131 244 7573
e-mail: shs@scotland.gsi.gov.uk

General enquiries on Scottish Government statistics can be addressed to:

Office of the Chief Statistician
Scottish Government
1N.04, St Andrews House
EDINBURGH EH1 3DG
Telephone: (0131) 244 0442
e-mail: statistics.enquiries@scotland.gsi.gov.uk

Further contact details, e-mail addresses and details of previous and forthcoming publications can be found on the Scottish Government Website at www.scotland.gov.uk/statistics

Complaints and suggestions
If you are not satisfied with our service, please write to the Chief Statistician, Mr Rob Wishart, 1N.04, St Andrews House, Edinburgh, EH1 3DG, Telephone: (0131) 244 0302, e-mail rob.wishart@scotland.gsi.gov.uk. We also welcome any comments or suggestions that would help us to improve our standards of service.

ScotStat
If you would like to be consulted about new or existing statistical collections or receive notification of forthcoming statistical publications, please register your interest on the Scottish Government ScotStat website at www.scotland.gov.uk/scotstat

Price Free