

Scottish Household Survey : Travel Diary

Background: Notes and Definitions

1. Introduction

A.1 This note accompanies statistics published on the Scottish Government website and provides detail information on the Travel Diary section of the Scottish Household Survey (SHS)

A.2 Note that detailed Travel Diary statistics are published every 2 years as 2 years of data are combined to produce detailed analysis. However annual headline results are released on alternate years: **2009 Travel Diary web tables were released on 22 November 2010.**

A.3 The SHS Travel Diary asks a random sample of adults (aged 16+) living in private households in Scotland about their travel the previous day. Statistics can be presented at a Scotland, local authority and Regional Transport Partnership area level, however 2 years of data are usually combined to present sub-Scotland analysis. Topics include:

- days of the week on which people travel,
- mode of transport used;
- purpose of journeys,
- distance/duration travelled,
- times of day at which trips start,
- characteristics of people making journeys
- number of car occupants,
- delays due to congestion to car journeys,
- delays to bus and rail journeys, whether a car driver paid for parking, and
- types of shopping journeys.

A.4 Prior to 2007 100% of households selected for the main SHS were included in the Travel Diary section. **From 2007 onwards this was reduced to 75%.**

Sampling Methodology and Weighting

A.5 The SHS is a continuous cross-sectional survey: interviewing takes place all year round. Each year, about 15,500 households across Scotland are interviewed. **75% of these are asked the Travel Diary.** The SHS is designed so that the interviews from each quarter will provide results which are representative of Scotland as a whole. In addition, the survey design is such that results will be available for each of the larger local authorities annually (those with an achieved sample of over 750 per year), and for all 32 Scottish local authorities, regardless of size, over two years. The SHS design therefore involves drawing a sample which will produce about 31,000 household interviews which are spread over two years. The requirement to produce results for every local authority, regardless of size, after two years, means that higher sampling fractions are used for Council areas with small populations, in order to ensure a minimum of about 550 household interviews in each area over the two years.

A.6 The results have been weighted to take account of differences in selection probabilities. As with all such surveys, sampling variability and non-response bias may affect the results. Results based on a small sub-sample size are more susceptible to this

and care should be taken when interpreting local authority and Regional Transport Partnership results. Figures for some local authority areas may be unrepresentative, because the interviewer(s) in those areas may conduct most of their interviews on certain days of the week. Although a reweighting process produces an even spread of interviews for Scotland as a whole, the distribution of reported journeys for some areas could depend greatly upon interviewing patterns as well as journeys made.

A.7 The weighting process ensures that the variation in sampling fractions does not make the results unrepresentative of Scotland as a whole. An average Council would have about 1,000 household interviews over the two years. Annual results are available only for the few Councils for which there are sufficient interviews per year. Of course, the smaller the sample upon which the results for a Council are based, the more carefully they will have to be used, and there might be cases where they do not appear sufficiently reliable to be used.

A.8 The SHS is intended to be a survey of private households. For the purposes of the survey, a household is defined as one person or a group of people living in accommodation as their only or main residence and *either* sharing at least one meal a day *or* sharing the living accommodation. A student's term-time address is taken as their main residence, in order that they are counted where they live for most of the year.

A.9 The sample was drawn from the Small User file of the Postcode Address File (PAF), which is a listing of all active address points maintained by the Post Office. The Small User file excludes addresses at which an average of more than 25 items of post are delivered per day. Blocks of flats etc, which have several dwellings at the same address, are *not* excluded from the Small User file: in such cases, the file's Multiple Occupancy Indicator is used to count each dwelling separately for the selection of the sample. Therefore, people in certain types of accommodation (such as nurses homes, student halls of residence, hostels for the homeless, other communal establishments, mobile homes, and sites for travelling people) will be excluded from the SHS unless the accommodation is listed on the Small User file of the PAF and it represents the sole or main residence of the people concerned.

A.10 In the areas of the nine Councils which have the highest population densities, the sample of addresses was drawn at random, within each geo-demographic stratum within each Council, at the start of the two-year period. The resulting addresses were then grouped into batches for allocation as interviewer assignments. For cost-effectiveness, the design of the sample clustered the interviews in the remaining 22 Councils. Enumeration Districts (EDs) were used as the Primary Sampling Unit. An ED contains, on average, about 150 households.

A.11 Totals may appear to differ slightly from the apparent sums of their component parts, in cases where they have been calculated by adding up the unrounded values of the components and then rounding each figure independently. Similarly, percentages may appear not to sum to 100 per cent.

A.12 The underlying sample numbers shown in different tables, or in different parts of the same tables may not be the same. This may be because some people were unable to, or did not want to, answer certain questions; the tables may relate to different populations (e.g. all adults, all journeys); for some cases, information is not available for some of the variables which were used to produce some of the tables, and therefore those cases cannot be included in some tables, or in sections of some tables.

A.13 The SHS Travel Diary collects information about travel for private purposes or for work or education, provided the main reason for the journey is not in the process of business. It includes the following types of travel - personal travel for domestic, social or recreational purposes and journeys made to take or escort someone else.

Journey Definitions

A.14 Journeys made by land, air or water within the United Kingdom are included. Journeys which start or end outwith the UK (e.g. a holiday flight from Spain) are excluded. However, if a respondent were to say that they had flown back from a holiday abroad on the previous day, the interviewer should record details of the journey home from the airport (but *not* record details of the flight to the UK).

A.15 The SHS Travel Diary does *not* cover: journeys which are made in the course of work by people who are employed as drivers or crew of public transport vehicles, to drive lorries, to deliver letters, parcels, leaflets or goods, as police officers etc. However, it does cover their journeys to and from their places of work; travel away from public roads or highways and recreational journeys.

A.16 The basic unit of travel, a journey, is defined as a one-way course of travel having a single main *purpose*. Outward and return halves of return journeys are treated as two separate journeys. If a single course of travel involves a mid-way change of purpose then it, too, is split into two journeys

A.17 **From 2007 Journeys less than ¼ mile or shorter than 5 minutes on foot are recorded.** Previously these were excluded. This is in an attempt to reduce any under reporting of short (likely to be) walking journeys. This has resulted in an increase in the proportion of walking journeys with corresponding decreases in the proportion of journeys by other modes. Care should be taken when comparing between years as some time series data not be directly comparable. Some time series data are less affected by this change (e.g. driver journeys delayed due to congestion).

A.18 A journey can consist of one or more stages. A new stage is defined when there is a change in the form of transport or when there is a change of vehicle requiring a separate ticket.

A.19 The **purpose** of a journey is normally taken to be the activity at the destination. Prior to 2007 a journey home was defined by the purpose at the origin of the journey (e.g. a journey from shops to home would be defined as shopping).

A.20 **From 2007 onwards** only a direct reverse journey of the outward journey (e.g. going straight home from work after travelling directly there earlier in the day) is classed as the origin's purpose (i.e. going to work). Non direct return journeys (e.g. going to the cinema before travelling home) would be defined by their own purpose (e.g. cinema, then going home). Hence from 2007 onwards a new category of "go home" exists (in addition to "go for a walk" resulting from the inclusion of short journeys under 5 min or ¼ mile – see A.15 above).

A.21 Some of the categories which are identified in the survey do not appear in subsequent tables presenting detailed analysis, as few journeys were recorded for them.

Mode of transport

A.22 Vans are included with cars; taxis and minicabs are in a separate category from ordinary cars; and there are separate categories for rail and underground, and for school

bus, works bus and ordinary (service) bus. However, some of these modes of transport do not appear separately in the tables, because few journeys were recorded for them. Therefore, the other category includes, motorcycles, ferries, aeroplanes and all other forms of transport that are not shown separately.

A.23 Where a journey involves more than one mode of transport (e.g. a bus then a train), the main mode of a journey is defined, the main mode of the journey is the one used for the longest (in distance) stage (as in the GB National Travel Survey (NTS)). This definition does *not* use the total of the distances travelled by each of the different modes to determine the main mode – e.g., a journey involving a 1 mile walk to a bus stop, a 1½ mile bus ride and a 1 mile walk to the ultimate destination is classified as ‘main mode = bus’, as bus is the mode of transport used for the longest stage of the journey, even though more than half the total distance is covered on foot. If there is no single longest stage, and the two (or more) longest stages do not involve the same mode of transport, the main mode of the journey is the mode used for the last of the longest stages. In practice, because of the way that the distances are calculated, it is unlikely that there will be many journeys which have two stages that involve *exactly* the same distance.

Highest Income Householder

A.24 The highest income householder is the household reference person for the first part of the main SHS interview. This must be a person in whose name the accommodation is owned or rented, or who is otherwise responsible for the accommodation. In households with joint householders, the person with the highest income is taken as the household reference person. If they have exactly the same income, the older is taken as the household reference person.

A.25 The Annual net household income is the total annual *net* income (i.e. after taxation and other deductions) from employment, benefits and other sources, which is brought into the household by the highest income householder and/or their spouse or partner. This includes any contribution to household finances made by other household members (e.g. for ‘digs’). Where household income is not present the SHS contractors have imputed the missing components of income, using information that was obtained from other households that appeared similar.

Day of the Week

A.26 The Travel Diary collects information about journeys that were made on the day *before* the interview: so, someone interviewed on Sunday will be asked about the journeys they made on Saturday. Journeys that start on one day and finish on another should be counted on the basis of the day on which they *started*.

A.27 Interviews are *not* spread evenly across the week, because some types of people are more likely to be found at home, available for interview, on certain days. Therefore, the results are weighted using factors, which depend upon the day of the week and the adult’s current situation (or economic status), so that, within each category of current situation, the *weighted* number of interviews are spread evenly across the days of the week. The weighting process covers *all* interviews, including those with people who had *not* made any journeys on the day before the interview. Therefore, the weighted numbers of people who said that they had made journeys, and the weighted numbers of journeys themselves, are *not* necessarily evenly spread over the days of the week.

A.28 Although the total number of weighted interviews are evenly spread across the week, this is not the case at the local authority level. Therefore, any analysis by day of week should be taken with caution.

Bias

A.29 The SHS results may be biased, tending to over-estimate the number of journeys, because the interviewer asks only about travel on the previous day: e.g. people may be more likely to be interviewed on the days on which they made no journeys than on the days on which they made many journeys, since they are more likely to be available for interview on days on which they have not made any journeys. Therefore, the probability of being interviewed on a particular day depends, to some extent, upon the amount of travel on that day. It follows that the day for which the information about journeys is collected (the day before the interview) does not represent a “completely random” choice of day, and therefore that the Travel Diary results may not be properly representative.

A.30 However, comparisons with (pre-2007) results of the GB National Travel Survey (NTS) suggest that the SHS Travel Diary *under*-estimates the number of journeys made by adults. This may have been because prior to 2007 journeys of less than a quarter of a mile, or of less than five minutes by foot were excluded. Also details of the previous day's travel are provided ‘off the top of the head’, as opposed to logged in a week long diary (as per the NTS) and therefore some journeys may be overlooked.

A.31 More detailed comparisons between the NTS and SHS Travel Diaries will be published early 2011. Detailed Scottish level information can be found at: <http://www.scotland.gov.uk/Topics/Statistics/Browse/Transport-Travel/PubTravScotRes>

Scottish Transport Statistics (due to be published 20 December 2010) will provide updated 2008/2009 NTS data in Chapter 12 *Personal and Cross-Modal Travel* and can be accessed at: <http://www.scotland.gov.uk/Topics/Statistics/Browse/Transport-Travel/PubScotTrans>

Imputation & Quality Assessment

A.32 Additional journeys have been imputed, in cases where it is obvious that they are missing – e.g. if the only journey recorded for the day was to work at 8.00 a.m., a return journey was imputed, using the same mode of transport and with the same duration. The imputation process uses information about the time spent at the destination by other people with the same current situation (economic status) who had reported making both an outward journey and a return journey for the same purpose. The average times spent at the destination, and the distributions of such times, are used to impute the times at which the return journeys would start. If the imputed time is after midnight, a return journey is not imputed.

A.33 Quality assurance procedures of Travel Diary data have also been improved, in light of the new Travel Diary structure. This has resulted some duplicate journeys deleted and some adjustment to raw data.

A.34 More information on the methods of imputation & quality assurance can be found in the Travel Diary User Guide, which is available on the SHS website: [Http://www.scotland.gov.uk/shs](http://www.scotland.gov.uk/shs)

Calculating Distance & Duration

A.35 The interviewer asks where the person started from, and where they went to, and records the origin and destination of each stage of each journey. When appropriate, the interviewer can specify that the previous destination is the origin of the current stage /

journey. Exact postcodes are determined/checked at a later stage in the processing of the data from the survey. In cases where only an approximate location is recorded (e.g. centre of Edinburgh), an arbitrary postcode (such as that of the main post office) is assigned. In some cases it may be unable to allocate a postcode from a postal district (e.g. EH10). Inevitably, there are occasions where no exact indication of location of the origin/destination can be determined. Continuous improvements to interviewers' computer systems result in improved location data over time..

A.36 The length of any journey stage is the estimated straight-line distance, based upon the grid co-ordinates of the centres of the postcodes of the origin/destination of that stage of the journey. In cases where the interviewer could not obtain sufficient details of the origin/destination to a postcode to be assigned, the Scottish Government had imputed the distance travelled. The distance of a multi-stage journey was calculated by adding up the distances of each of its component stages. For series of calls journeys, the respondent estimates the total distance for series of calls journeys.

A.37 As most journeys are not made in a straight line, the distance will underestimate the actual distance travelled.

A.38 Prior to 2007 the duration of a journey was calculated from the start and end times. As the recording process will only be accurate to - at best - say the nearest five minutes. the estimated durations of some journeys would be subject to possibly large percentage errors. Due to coding problems in the CAPI script in October, November and December 1999, the start time and end time of some journeys are missing for around 4 per cent of journeys for 1999 as a whole. As duration is derived from the start time and end time of journeys, about 7 per cent of journeys in 1999 have a missing duration.

A.39 From 2007 onwards duration is collected direct from the respondent. This aims to improve the accuracy of the data. This means that data prior to 2007 may not be strictly comparable.

Urban/Rural Classification

A.40 The urban/rural classification shown in the tables was developed for use in analysing the results of the SHS. It is based on settlement sizes and for the less-populated areas the estimated time that would be taken to drive to a settlement with a population of 10,000 or more. The classification is based on postcodes. Six categories have been defined:

- **Large urban areas** - settlements with populations of 125,000 or more.
- **Other urban areas** - other settlements with a population of 10,000 or more.
- **Accessible small towns** - settlements of between 3,000 and 9,999 people, which are within 30 minutes drive of a settlement of 10,000 people or more.
- **Remote small towns** - settlements of between 3,000 and 9,999 people, which are *not* within 30 minutes drive of a settlement of 10,000 people or more.
- **Accessible rural areas** - settlements of fewer than 3,000 people, which are within 30 minutes drive of a settlement of 10,000 people or more and
- **Remote rural areas** - settlements of fewer than 3,000 people, which are *not* within 30 minutes drive of a settlement of 10,000 people or more.

Deprivation

A.41 The Scottish Index of Multiple Deprivation (SIMD) was designed by the Scottish Government to rank the data zones used for the production of Scottish Neighbourhood Statistics in order of deprivation. There are 6,505 data zones, with an average of about 750 residents in each, formed by aggregating Census output areas. The SIMD is based on 38 indicators in the seven individual domains of Current Income, Employment, Health, Education, Skills and Training, Geographic Access to Services, Crime and Housing. More information can be found at the SIMD website:

(<http://www.scotland.gov.uk/Topics/Statistics/SIMD>).

A.42 Households in the SHS sample have been allocated the SIMD value of the data zone which contains the postcode of the residence. In the small number of cases where a postcode is split between more than one data zone, the SIMD value used is that of the data zone into which the largest number of dwellings in that postcode falls. The SIMD values have further been assigned to one of 5 quintiles, with quintile 1 containing the most deprived 20 per cent of data zones in Scotland, and quintile 5 the least deprived 20 per cent. Because the SHS sample is not spread uniformly across Scotland, the quintiles do not necessarily each contain exactly 20 per cent of the households in the SHS sample.

Local Area Results

A.43 The Regional Transport Partnership (RTP) areas and the local authority areas that they cover are shown in the below. In the case of an RTP which covers the area of only one local authority (or of only a few local authorities), the figures may be subject to quite large sampling errors if the SHS does not have a large sample for that area.

<i>Regional Transport Partnership</i>	<i>Local authority area(s) included:</i>
Central & Tay	Angus, Dundee City, Perth & Kinross, Stirling.
Highlands & Islands	Argyll & Bute (<i>except Helensburgh and Lomond</i>), Eilean Siar, Highland, Moray, Orkney.
North East	Aberdeen City, Aberdeenshire.
Shetland	Shetland.
South East	Clackmannanshire, East Lothian, Edinburgh City, Falkirk, Fife, Midlothian, Scottish Borders, West Lothian.
South West	Dumfries & Galloway.
Strathclyde	East Ayrshire, East Dunbartonshire, East Renfrewshire, Glasgow City, Inverclyde, North Ayrshire, North Lanarkshire, South Lanarkshire, Renfrewshire, South Ayrshire, West Dunbartonshire and the Helensburgh and Lomond areas of Argyll & Bute.

A.44 The figures for the Highlands & Islands include the *whole* of the Argyll & Bute Council area (including the Helensburgh and Lomond areas), and the figures for Strathclyde do *not* include *any* of Argyll & Bute. As Helensburgh and Lomond contain much less than half of the total population of Argyll & Bute, their being counted in the wrong RTP area for the purpose of preparing the figures reported in this bulletin is unlikely to affect greatly the results.

A.45 Although the SHS's sample is chosen at random, the people who take part in the survey will not necessarily be a representative cross-section. Results produced from a small sample could be affected by sampling variability. The larger the sample, the less likely it is that the results will be affected greatly by sampling variability. Thus caution must be used when analysing results from small sub samples.

Sampling Variability and Confidence Limits

A.46 The likely extent of sampling variability can be quantified, by calculating the standard error associated with the estimate of a quantity produced from a random sample. Statistical sampling theory states that, on average only about one sample in three would produce an estimate that differed from the (unknown) true value of that quantity by more than one standard error; only about one sample in twenty would produce an estimate that differed from the true value by more than two standard errors; only about one sample in 400 would produce an estimate that differed from the true value by more than three standard errors. By convention, the 95 per cent confidence interval for a quantity is defined as the estimate plus or minus about twice the standard error (from sampling theory, the interval is plus or minus 1.96 times the standard error), because there is only a 5 per cent chance (on average) that a sample would produce an estimate that differs from the true value of that quantity by more than this amount.

A.47 Table 13 shows the 95 per cent confidence limits for estimates of a range of percentages calculated from sub-samples of a range of sizes (NB: the confidence limits for estimates of x per cent and for $(100-x)$ per cent are the same). The table was produced in the same way as the tables of 95 per cent confidence limits in the Annual Report volumes of *Scotland's People*, but has a more detailed breakdown of the smaller sample sizes.

A.48 The interpretation of an entry in Table 13 is best explained by an example:

- The value in the cell at the intersection of the 45 per cent or 55 per cent column and the 800 row is 4.5;
- this means that the 95 per cent confidence limits for an estimate of 55 per cent which is produced from a sub-sample of 800 are ± 4.5 percentage-points;
- so the 95 per cent confidence interval for the estimate is 55 per cent ± 4.5 percentage-points (i.e. from about 50.5 per cent to around 59.5 per cent, assuming that the value of the estimate is 55.0 per cent).

A.49 As the survey's estimates may be affected by sampling errors, apparent differences of a few percentage points between the figures for two sub-groups of the population may not be significant: it could be that the true values for the two sub-groups are similar, but the random selection of households for the survey has, by chance, produced a sample which gives a high estimate for one sub-group and a low estimate for the other. One way of assessing significance at the 5 per cent level involves comparing the difference with the 95 per cent confidence limits for the two estimates.

- Suppose that these are ± 3.0 percentage-points and ± 4.0 percentage-points, respectively.
- Clearly a difference which is *less* than the magnitude of the largest limit (4.0 percentage-points) is *not* significant; and a difference which is *greater* than the *sum* of the magnitudes of the limits (3.0 percentage-points + 4.0 percentage-points = 7.0 percentage-points) is significant.
- Statistical sampling theory suggests that a difference whose magnitude is between these values is significant *if* it is greater than the square root of the sum of the squares of the magnitudes of the limits for the two estimates - in this case, $(3.0^2 + 4.0^2)^{0.5} = 5.0$.

- So a 5.0 percentage-point difference would be considered statistically significant (at the conventional 5 per cent level). However, one may well find some apparently significant results that are actually just the result of sampling variability, having arisen by chance.

A.50 The above information relates only to sampling variability. The survey's results could also be affected by non-contact/non-response bias: the characteristics of the people who should have been in the survey but who could not be contacted, or who refused to take part, could differ markedly from those of the people who were interviewed. If that is the case, the SHS results will not be representative of the whole population. Without knowing the true values (for the population as a whole) of some quantities, one cannot be sure about the extent of any such biases in the SHS. However, comparison of SHS results with information from other sources suggests that they are broadly representative of the overall Scottish population, and therefore that any non-contact or non-response biases are not large overall. The *Fieldwork Outcomes* and *Methodology* volumes of *Scotland's People* provide more information on these matters.

A.51 Figures are rounded to 1 decimal place. Aggregating sub categories (e.g. bus and rail to obtain a public transport figure) may result in slightly inaccurate figures due to this rounding.

A.52 Anonymised copies of the survey data are deposited at the UK Data Archive.

Enquiries about Travel Diary statistics in this bulletin should be addressed to:

Transport Statistics
Transport Scotland
Scottish Government
Victoria Quay
Edinburgh, EH6 6QQ
Tel: 0131 244 1457
Fax: 0131 244 0888
E-mail: transtat@transportscotland.gsi.gov.uk

General enquiries about the SHS should be addressed to the survey's Project Manager:

SHS Project Manager
Social Research
Scottish Government
Victoria Quay
Edinburgh, EH6 6QQ
Tel: 0131 244 8420
Fax: 0131 244 7573
E-mail: shs@scotland.gsi.gov.uk

Please inform the Project Manager if you wish to be added receive *e-mail* updates of SHS developments. Those without email/internet access can request postal updates.