Incorporating survey data into household projections

1. Purpose

The General Register Office for Scotland (GROS) are investigating survey data to determine whether this data can be used to update information on headship rates that are used in household projections.

Note

Household Analysis Review Group (HARG) members are asked to consider whether it is appropriate to further investigate using survey data for future household projections and make suggestions for the form this may take.

2. Background

Household projections are based on the most recent population projections produced by the GROS. The number of people living in private households is estimated by taking the population projections for each year and subtracting the number of people living in communal establishments, such as student halls of residence, care homes or prisons.

To estimate the number of households of each type, information on household type and age group is projected forward from the 1991 and 2001 Censuses, for each household type, age group and local authority area. This information is applied to the private household population to produce the basic household projections. Because the overall projections for Scotland are believed to be more accurate than those for individual local authorities, the local authority figures are constrained to the Scottish total.

Each year GROS produces household estimates based on Council Tax data. These estimates do not include information on household type or age group, but they do provide an overall estimate of the number of households in each local authority area, based on more up-to-date data than the household projections. Therefore household projections are also adjusted to match the household estimates.

A weakness with this approach is that it is reliant on Census data, which is only available every ten years. For the 2006-based household projections we projected figures up to 2031, based on information on the composition of households from the 1991 and 2001 Censuses. We used the 2006-based population projections and we constrained to the latest household estimates but this is an overall figure of the total number of households in each Local Authority, and does not include any information on household type. Trends in household formation may have changed since the last Census. Our current methodology will not reflect these changes. It may improve the projections if we are able to incorporate more up-to-date information on household type between Censuses.

3. Survey Data

The most likely option for updating the household information between Censuses is to use surveys. Unlike the Census, surveys only cover a sample of the population

but they are carried out every year (or on a rolling basis), so they can provide more up-to-date overall figures. As a minimum they may ask about the number of people in the household by age and gender which is all the information we need to determine household type. The following are a few surveys which were considered.

3.1 The Scottish Household Survey (SHS)

The SHS covers around 15,000 households per year, approximately 0.7% of households in Scotland. The structure of the survey is a continuous cross-sectional survey, each complete sample being covered in the course of two years. The sample is drawn from the small user file of the Postcode Address File (PAF). The overall design of the sample is to pursue a wholly random sample where fieldwork conditions allow, namely in areas of high population and to cluster interviews in the remaining areas (also on a random basis). Data that are currently available is for the years 1991 to 2007

3.2 Labour Force (LFS) and Annual Population Surveys (APS)

The APS is a UK-wide survey. It is a modular survey, which incorporates what was known as the Labour Force Survey (LFS). From March 2003, the APS sample in Scotland was boosted from 8,000 households to around 23,000. The APS data are based on calendar quarters and are published quarterly on a rolling annual basis.
Data are currently available from the LFS for years 1999–2003 and then from the APS for years 2004–2008.

3.3 Scottish Crime and Justice Survey (SCJS)

The Scottish Crime and Justice Survey is a social survey which asks people about their experiences and perceptions of crime in Scotland. Respondents are selected at random from the Postal Address File and participation in the survey is entirely voluntary. The survey involves interviewing a randomly selected adult in 16,000 households across Scotland per year.
Data are currently available for the years 2006 and 2008/9

4. Weighting

Unlike the Census, surveys are voluntary so they are affected by under- and over-counting. For instance, young men are less likely to be included and older couples are more likely to be included.
SHS data provided to us (and published by the SHS team) is not adjusted to account for this. Therefore SHS figures have been modified by comparing the 2001 SHS figures with the 2001 Census and adjusting accordingly.
Unlike the SHS, APS results are adjusted to account for under- and over-counting. Data are calibrated to correct for non-response at local authority level, and by age group and gender.
The SCJS uses two weights - individual and household weights. Individual weights are used to gross up to the total population in Scotland and household weights are used to gross up to the total number of households in Scotland. This results in constant proportions for each of the household types for the 2 years of data provided therefore it seems that we will not be able to use SCJS data

5.   Comparison of Projected Headship Rates with Survey Data

The charts shown in Annex 1 compare proportions from projected headship rates and survey data for different household types. The graphs for projections are from household projections publications and are derived from 1991 and 2001 census data and therefore based on the entire population (2,200,000 households). Survey data is based on a small sample of the population and is therefore likely to be more variable. For the SHS the sample size is around 15,000, for the LFS the sample size was 8,000 and for the APS sample size is approximately 23,000. SHS data have been plotted from 1999 to 2007, LFS data have been plotted from 1999 to 2003 and APS data have been plotted from 2004 to 2008.

As may be expected survey data proportions are quite variable and have a fairly large associated confidence interval. It can also be seen that trends seem to differ between sets of survey data.

6.   Improving Survey data

The reliability of the SHS and APS survey data may be improved if results from each are combined. However, as the samples for these surveys are drawn in different ways it is not clear whether this would be feasible. This would have to be investigated more thoroughly.

The survey results fluctuate across years so it may be worthwhile using some method of smoothing. This would be more appropriate for data from the SHS as we have data from 1991 to 2007. This may not be useful for the LFS and APS which cover shorter time periods.

It may be possible to combine survey data with the census projections for headship rates. In this case it would be sensible to weight the data sources in some way. Census data should perhaps have a large weight because it is based on the entire population. However, survey data may also command a large weight because it is more recent. Deciding on an appropriate weighting strategy will require some consideration.

7.   Next Steps

There could be an advantage to using survey data if trends in household formation have changed since the last Census and if we can incorporate more up to date information on household type then this would improve our projections.

For 2010-based projections we intend to use information from the microsimulation model (SCOTSIM) created by Ashley McCormick from Liverpool University to improve our household projections. This microsimulation model has been created using the British Household Panel Survey and will provide more up to date information on household types.

In preparation for incorporating this new model we plan to evaluate current 2008-based household projections against SHS, APS survey data and SCOTSIM results.

Note: HARG members are asked to consider whether it is appropriate to further investigate using survey data for future household projections and make suggestions for the form this may take.

GROS: Household estimates branch
March 2010

Annex 1: Charts comparing proportions in varying household types

1 adult: male

Legend: Projections, SHS, LFS, APS

1 adult: female

Legend: Projections, SHS, LFS, APS

2 adults

1 adult, 1 child

1 adult, 2+ children

Legend: Projections, SHS, LFS, APS

2+ adults, 1+ children

Legend: Projections, SHS, LFS, APS

3+ adults

Legend: Projections, SHS, LFS, APS