Household Analysis Review Group (HARG)

Household Projections for Scotland: Methodology Proposals following the 2011 Census

Contents

1. Purpose

To look at the methodologies used by other countries in the production of household projections as we consider whether there should be any changes made to the methodology of our household projections after the release of 2011 Census data.

The countries that have been looked at are England, Wales, Northern Ireland, France, the Netherlands, Luxembourg, South Korea, Japan, Hong Kong, Canada, Australia and New Zealand.

2. Existing methodology

2.1 Data

The household projections are based on the sub-national population projections that are also produced by the National Records of Scotland (NRS).

A range of administrative data sources are used to calculate the number of people living in communal establishments, which is subtracted from the total population to estimate the number of people living in private households.

Information on the types of households, in the form of headship rates, is taken from 1991 and 2001 Censuses. The two most recent years' household estimates are used to constrain the household projections.

2.2 Method

The proportion of individuals of a particular age and gender within each local authority area who head households is called the headship rate, where the 'head' of the household is normally the first adult entered on a census form. The non-headship rate is the proportion of people who are not head of households within each age group, gender and local authority area.

These headship and non-headship rates from the 1991 and 2001 censuses are fitted to a two point exponential curve and projected forward across the 25 year projection period. Applying these rates to the projected private household population gives the number of households by type of household, age of the head of household and local authority. Local authority figures are then constrained to the Scottish household projections and the two most recent household estimates and adjusted to ensure that there are enough people in the population projections to fill all the households.

The current household types and age groups that we use are shown in Tables 1a and 1b.

Note:
We would welcome comments from HARG members regarding whether the household types and age groups meet user requirements.

Tables 1a and 1b: The household types and age groups that are used in the current household projections for Scotland

Household Types
1 Person Households:
• 1 adult: male
• 1 adult: female
2 Person Households:
• 2 adults
• 1 adult, 1 child
3+ Person Households:
• 1 adult, 2+ children
• 2+ adults, 1+ children
• 3+ adults

Age Groups
16-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
85-89
90+

3. Other Methodologies

We looked at the household projections for 12 other countries; a brief summary of the methodology used in each country can be found in Annex A. Two of the most important parts of the methodology for household projections are the choice of household formation information and how this is projected. Table 2 provides a summary of these two pieces of information for each country.

3.1 Household Composition Information

The use of headship rates, our current method, is one of the most common methods used in producing household projections among the countries that we looked at. The popularity of this method is due to it being simple to implement and having relatively modest data requirements. The definition used for the head of a household varied among the countries, as did the method of projecting the headship rates.

Membership rates are also a popular choice among the countries we looked at. Membership rates are the probability of someone in a particular age and gender group being part of a particular household type, this means there is information on all members of households rather than just the head of the household. This means that age breakdown of the households is quite different from household projections that use a headship method, see Annex B.

Finally, the Netherlands and Japan use more dynamic models in the production of their national household projections. These models are much more complex than using headship or membership rates as they relate household formation directly to the probabilities for individuals in a particular age and gender group experiencing events such as marriage or having a child. Due to the increased complexity of models such as these the data demands are much greater. Dynamic models are not considered any further in this paper as we are currently waiting for a handover of SCOTSIM, which is a dynamic micro-simulation model for Scotland (a separate oral update will be provided on this).

Table 2: A summary of the methodologies used in other countries for household projections

	Household composition information	Projection method
Canada	Headship	Assumes a decline/increase in headship rates for certain age groups but doesn't say how this is calculated.
England	Headship	Stage 1: weighted between dampened logistic trend and simple logistic trend Stage 2: two point exponential
Scotland	Headship	Two-point exponential
Hong Kong	Headship	'Life cycle model'
Japan (sub-national)	Headship	Uses the relative difference between prefecture headship rates and national headship rates
Japan (national)	Dynamic Transition Model	n/a
South Korea	Headship	Log-linear formula and net transition rate
France	Membership/Headship	Quasi-linear
Australia	Membership	Linear
New Zealand	Membership	Linear/Constant
Luxembourg	Membership	Two-point exponential
Northern Ireland	Membership	Two-point exponential
Wales	Membership	Two-point exponential
Netherlands	Macro-dynamic model (LIPRO)	n/a

3.2 Method of Projection

For the countries that use headship rates or membership rates, a range of methods are used to project the rates. Often the choice of projection method appears to have been largely influenced by the number of data sources that were used to look at the trends in the headship/membership rates.

France, Australia and New Zealand use data from more than two censuses when looking at the trends in headship/membership rates; due to the number of data points, these countries all use reasonably simple projection methods that are variations of linear projections.

Meanwhile Wales, Northern Ireland and Luxembourg are countries that only use two censuses and all use the same modified two-point exponential formula that we currently use.

The use of more basic projection methods when more than 2 data points are used is highlighted by the methodology used in the English household projections. The logistic trends used in the first stage of their projections incorporate four censuses worth of data as well as data from the Labour Force Survey. This is easier to apply than an adapted version of the two-point modified exponential formula used in the second stage where just two censuses are used.

4. Options

This section outlines the major choices that have to be made regarding the household projections methodology following the release of the 2011 Census.

4.1 How many censuses should we use

In the last set of household projections we used data from the two most recent censuses. With the release of the 2011 Census we will have the choice of using two or three censuses, this choice will be related to the choice of household composition information as it would not be possible to use the 1991 census for all of the options.

4.1.1 Option 1 – Use the 2001 and 2011 Censuses

Advantages:

- Only using two data points would mean that it would still be possible to use the modified two-point exponential formula.

- It would give the option of using the household reference person, discussed in Section 4.2.2.

- It will make it simpler for users who want to replicate / modify the NRS projections, for example by using software such as the household projections module in POPGROUP.

- This effectively means there would be no change from the current methodology; we would just update the figures with more recent data. Therefore, this is the simplest approach, and probably the easiest for users to understand, and to make comparisons with previous projections.

Disadvantages:

- The trends in this period will be affected by the global economic downturn as this has had an impact on new household formation, for example "young adults – particularly men – who are unemployed have become increasingly likely to be living in the parental home during their early twenties." If this impact is temporary then using these Censuses would lead to an underestimate in the number of households once household formation returns to the levels observed before the economic downturn.

4.1.2 Option 2 – Use the 1991 and 2011 Censuses

Advantages:

- Only using two data points would mean that it would still be possible to use the modified two-point exponential formula.

- It will make it simpler for users who want to replicate / modify the NRS projections, for example by using software such as the household projections module in POPGROUP.

- By using censuses that are further apart, the effect of the global economic downturn on the trends in household formation will be reduced. The result would be that projected household formation rates should be somewhere

between the rates witnessed in the 90s and the reduced rates during the economic downturn.

Disadvantages:

- Information from the 1991 Census may not be relevant to future changes in household formation rates as a reasonable amount of time has passed since then.

4.1.3 Option 3 – Use the 1991, 2001 and 2011 Censuses

Advantages:

- By using trends from before the global economic downturn as well as ones witnessed during it, the result may be a trend that better reflects future trends in household formation once the impact of the economic downturn has reduced. This would only be an advantage if the change in household formation rates caused by the economic downturn is thought to be temporary.

- We would be able to choose a projection method that provides flexibility to build in changes in trends by giving the data from different censuses different weights.

Disadvantages:

- Information from the 1991 Census may not be relevant to future changes in household formation rates as a reasonable amount of time has passed since then.

- A different method of projection would be required as the modified two-point exponential formula would not be suitable. There is more information on this in Section 4.3.

4.2 Household composition information

There are three options for what household composition information we use. The first is to continue to use headship rates in exactly the same manner as we currently do with the head of household defined as the first adult on a census form. The second is to use headship rates but to use the household reference person as the head of household, the household reference person incorporates economic status and age into the choice of the head of household. The third option is to use membership rates.

4.2.1 Option 1 – Headship rate using current definitions

Advantages:

- As this method is already in use, new sets of household projections that are produced will be comparable with the old ones.

- All of the options regarding the number of censuses we use would be suitable if we used this.

Disadvantages:

- This method is sensitive to who completes the census form.

- This method is more sensitive to changes in the older male population than sub-populations as traditionally this is who would be entered first on a census form. However this bias is probably not as strong as it was in the past.

4.2.2 Option 2 – Headship rate using household reference person

This option was considered in 2002 following the release of the 2001 Census and produced similar results to the projections produced using the current methodology. The current method was preferred as projections using the household reference person showed more volatility in some of the smaller Local Authorities due to data on the household reference person only being available for a 10% sample of the 1991 Census.

Advantages:

- By having a well defined set of rules regarding economic status and age to choose the head of household it will not be affected by changes in who completes the census form in a household.

- As this method produces similar results to projections produced using the current method any new projections would be reasonably comparable to older sets of projections.

Disadvantages:

- It would not be possible to use the 1991 Census as there is only a 10% sample with this information, so we would have to use the 2001 and 2011 Censuses.

4.2.3 Option 3 – Membership Rates

Membership rates are the probability of someone in a particular age and gender group being part of a particular household type. As this method uses information on all household members rather than focussing on the head of a household, the outputs using this method would be different from those that are currently produced. Using this method would mean the information available regarding the age structure would change. While we would no longer be able to provide information on the age of the head of the household (apart from for households with one adult), the proportion of each age/gender group living in each household type would be available, see Annex B for a more detailed summary of the age/gender breakdown that would be possible.

Advantages:

- This method avoids bias towards any age or gender group as all household members are considered rather than only considering the head of the household.

- It would be possible to provide the projected proportions of each age/gender group that live in a particular type of household.

Disadvantages:

- The new sets of household projections that would be produced would not be easily comparable to older projections due to the methodology change.

- It would take more work to change to this methodology than continuing with a headship method as the systems are already in place for a headship method.

- The number of household types would have to increase to include households of larger sizes to be able to calculate good household number and average household sizes. This would increase the complexity of the data required to produce the projections and mean that projections will be being produced based on population groups that are very small.

4.3 How should the relevant rates be projected

The options for this will depend on the previous choices. If only two censuses are used then we will be able continue using the modified two-point exponential formula that we currently use.

If all three censuses are used then it will be necessary to find a new method of projection. One option is use a 'weighted two-point exponential method' to project the rates. This would use the two-point exponential formula for each pair of censuses separately to produce three sets of rates; weights could then be applied to these projected rates to produce a set of projected rates that is a based on a mixture of the trends seen between 1991-2001, 1991-2011 and 2001-2011. The weights could be determined by calibrating to an external data source such as Scottish Household Survey household type proportions or similar data from the new core harmonised question dataset. This method of projection would also have the advantage that it would make it easy to produce variant projections by changing the weights.

Another option is to use a three-point modified exponential model, however this is considerably more complex than the two-point version. This method was considered in the 1990s, however the projected households with children from this method did not follow the pattern expected when compared to the projected numbers of children aged 0 to 15 from NRS so the two-point exponential method was preferred.

If all three censuses are used we would view the weighted two-point exponential method of projection as the preferred option. It would be easier to program than the three-point exponential method, it would be easier to produce variant projections and it would be more easily replicated by users using POPGROUP or similar software.

5. Conclusion

Table 3 gives a summary of the methods that we consider as being the most likely to be chosen.

Table 3: Summary of the preferred options

Household Composition Information	Censuses Used	Projection Method
Headship – First name on census form	1991, 2001 and 2011	Weighted two-point exponential projections
	1991 and 2011	Two-point exponential
Headship – Household reference person	2001 and 2011	Two-point exponential

We don't have to make a decision just yet on which method to use. Once we have access to more detailed information from the 2011 Census, we can try producing projections using one or more approaches, and we can then discuss the results with HARG, before making a decision together in time for the 2012-based projections due out in Summer 2014. We could use the existing 2010-based population projections for this analysis, as the 2012-based projections will not be available until Spring 2013. However, it would help a lot if we could at least narrow down the list of options we want to consider, so we would appreciate HARG members' views on this. To help with this, we have set out our initial proposals, below.

We feel that the use of membership rates would not have sufficient advantages over the use of headship rates to justify the amount of work that would be required to change to this methodology. There would also be a loss of comparability with previous household projections. Therefore we would recommend that we continue to produce household projections using a headship method.

Only using the 2001 and 2011 Censuses in the household projections has a significant disadvantage due to the effect of the global economic downturn on the trends in household formation. For this reason we would prefer to continue using the head of household rather than the household reference person, as this would give us the option of using the 1991 Census to reduce the impact of the economic downturn.

Using all three censuses with the weighted two-point exponential method of projecting headship rates would probably be our preferred option. By using the 1991 Census we would be reducing the impact of the global economic downturn, while the weighted method of projection would build in the flexibility to adapt the projections to future changes in trends, particularly as we get further from the 2011 census year, and an easy method of producing variants.

Note:
We propose that the preferred method of producing the household projections is to continue using the head of household as defined in previous years. However we would use the 1991, 2001 and 2011 censuses and project the headship rates using the weighted two-point exponential method. We would welcome the views of HARG members' on this. We have scope to try out different methods and discuss the results with HARG members before producing the 2012-based projections next year.

6.   Next Steps

- Request the relevant data from the 2011 Census.

- Modify the computer programs used, to allow us to produce the projections using the methodology (or methodologies) agreed with HARG members.

- Produce projections using the 2010-based population projections, for now.

- Analyse and discuss the results with HARG members, to agree the method.

- we will use for the 2012-based household projections.

NRS: Household estimates and projections branch
April 2013

Annex A – Summaries of methods used by other countries

England

England, like Scotland, uses headship rates although the methodology used is more complicated than ours as it is divided into two stages.

The first stage projects household representative rates that are obtained from the 1971, 1981, 1991 and 2001 Censuses as well as Labour Force Survey (LFS) using a combination of a simple logistic trend and a dampened logistic trend. These projected rates are then applied to the private household population to obtain a figure for the overall number of households.

The second stage is to use 1991 and 2001 Census commissioned tables to disaggregate the overall number of households from stage one into more detailed household types. This is done by obtaining headship rates from 1991 and 2001 Census tables and projecting them forward using the same modified two-point exponential model that we currently use. The projected headship rates are then applied to the private household population to give the number of households of each type that are headed by someone from each age group, these figures are then constrained to the overall number of households that were calculated in stage one.

Wales

The Welsh household projections are produced using household membership rates rather than headship rates. A household membership rate is the proportion of each age and gender group that live in a particular household type. This means that the household types that are produced are different from what we currently use as there is information on all of the household members, however there is no information about the age of the head of household.

Household membership rates are calculated from the 1991 and 2001 censuses and projected forward using a modified two-point exponential model. These rates are applied to the private household population to find the size of the population that live in each type of household which then allows the number of households of each type to be calculated by dividing the population by the number of people in that household type.

Northern Ireland

The methodology used by Northern Ireland is similar to the methodology used by Wales with some small differences, such as the household membership rates for persons under 18 being kept constant at 2001 Census levels as trends that were observed between 1991 and 2001 appear not to have continued based on more recent demographic evidence.

Luxembourg

The methodology used for household projections for Luxembourg is based on the methodology used in the household projections for Northern Ireland and is therefore only very slightly different.

France

The national household projections for metropolitan France are based on the national population projections by single year of age and gender and use household membership rates. First of all household membership rates are projected based on data between 1990 and 2005 using a "quasi-linear" method, where weightings of the trends between 1990 and 1999 (based on a traditional census) and the trends between 1999 and 2005 (the first rolling census period) are used.

The household members are then converted into households using household representative rates as follows:

- one for people living alone (by definition)

- zero for women in couples or for children (the French definition for household representative is the oldest economically active male)

- fixed at 2005 levels for other groups.

The sub-national projections use a headship based method and are based on the sub-national population projections and projections of the household representative rates for Metropolitan France and the household representative rate for region concerned from the last census (the base year of the Metropolitan France projection). These rates are split by age and gender, but household type is not taken into account. The household representative rate for the region is projected using the following formula:

$$T_t = \frac{T_{1999} \cdot TF_t \cdot (1 - TF_{1999})}{T_{1999} \cdot TF_t \cdot (1 - TF_{1999}) + TF_{1999} \cdot (1 - TF_t) \cdot (1 - T_{1999})}$$

Where T1999   = household representative rate for 1999 for the region
           TF1999 = household representative rate for 1999 for Metropolitan France
           Tt   = projected household representative rate for year t for the region
           TFt   = projected household representative rate for year t for Metropolitan France

These household representative rates are applied to the population projections to estimate household numbers.

The Netherlands

The method used by Statistics Netherlands until 2009 is much more complex than the methods that have already been discussed, consisting of four separate models.

The first step involves projecting the population by marital status. This is computed using a macro-simulation model which generates marital status transitions, deaths and births in the start of year population and applies net migration figures.

The next step is to use a static macro model to project the population by household position and marital status. Essentially this model distributes the projected population by age, gender and marital status over the household positions. Once this has been done all persons with the household position 'single', 'single parent' or 'head of household of type other' are reference persons in their household. Women with the household position 'living with a partner' are also considered to be reference persons. The number of households is then calculated from the number of reference persons.

The third step uses a dynamic macro model to project household composition by size and number of dependent children. This step models changes in household composition due to the birth of children and children leaving the parental home.

Finally the number of first and second generation immigrants is projected, this projection is based on assumptions about how the age and gender distributions of the household positions for immigrant groups will change compared to those for the total population

Since 2009 the household projections for the Netherlands have been produced using LIPRO, a multidimensional demographic projection software produced by the Netherlands Interdisciplinary Demographic Institute.

Canada

Canada is another country that uses a headship method to produce household projections, where the head of a household is defined as the person or one of the people who are responsible for the major costs of maintaining a private household. Headship rates are obtained from census data for heads of family households and non-family households. The headship rates are projected in three ways, a low projection, a medium projection and a high projection. The low headship rate scenario assumes a general decline in headship rates across all age groups over the projection period. In the high headship rate scenario, headship rates are projected to rise among young and middle-aged households, as well as among older households up to the 60-64 age group; the headship rates are assumed to decline, however, among those in the 65 and older age category. The medium scenario represents an average of the age-specific headship rates in each of the high and low scenarios.

The projected headship rates are applied to the projected population by age and gender to produce figures for the total number of family households and non-family households.

House ownership rates are then projected forward, either by keeping 2006 Census levels constant or continue the trend in age-specific ownership rates observed between 1996 and 2006 but at a decreasing rate. Rates of the types of dwellings that people live in are also projected forward by keeping them constant at 2006 Census levels. Both of these rates are then applied to the total number of family and non-family households to give the number of family and non-family households, by tenure, dwelling type & age.

Hong Kong

Hong Kong uses headship rates when producing household projections. Observed headship rates are obtained from censuses and by-censuses (one of which is held every 5 years) from 1981 to 2011. The headship rates are projected forward so that they vary smoothly for each cohort using a "life cycle" model.

The projected headship rates are then applied to the projected population to give the number of households by size and type for the next 10 years while also giving the total number of households for a further 20 years.

South Korea

South Korea is another country that uses headship rates, they considered membership rates but it was felt that it was unclear whether the extra information gained would merit the increased difficulties in applying this method.

Marital status proportions are projected using the following log-linear formula and then applied to the projected population.

Headship rates are projected forward in two ways the log-linear formula for the under 35 and net transition rates for those aged 35 and over.

The log-linear formula:

$$C^s(m,x,t) = \frac{H^s(m,x+5,t+5)}{H^s(m,x,t)} - \frac{P^s(m,x+5,t+5)}{P^s(m,x,t)}$$

where   Cs(m,x,t) is the net transition rate in sex s, age x, marital status m and year t
        Hs(m,x,t) is the number of households in sex s, age x, marital status m and in year t
        Ps(m,x,t) is the number of people of sex s, age x, marital status m and in year t

The projected headship rates are then applied to the projected population to give the projected number of households by sex, age and marital status.

The projected number of households by composition and household size are also produced using projected distributions of these features.

Japan (National)

The national household projections for Japan are produced using a dynamic 'household transition' model. This model is based on the probabilities of changes in an individual's position in a household, these are also known as transition rates. Probabilities are calculated and projected by age for household transitions such as leaving parental home, first marriage, remarriage, divorce, death of spouse, etc. and these are used to inform changes in household composition which can then be used in the calculation of household projections.

This method relates household formation and dissolution directly to demographic determinants and is much more dynamic than the methods using headship/membership rates. However the data requirements for this type of model are much greater than using headship rates.

Japan (Sub-national)

Japan uses a different methodology for sub-national household projections, preferring a headship based method for the 47 prefectures of Japan although it does incorporate some information from the national household projections.

The first step in projecting the headship rates for each prefecture is to calculate the national headship rates by age, gender and family type for each year in the projection period from the national household projections.

Past trends of the relative difference between prefecture headship rates and national headship rates are then analysed with the data coming from the 1980, 1985, 1990, 1995, 2000 and 2005 Censuses. If the relative difference has been increasing constantly in the past then it is assumed that the relative difference will remain constant at the 2005 level, otherwise it is assumed that the difference will decrease over time in the projection period.

The projected prefecture headship rates are then applied to the projected prefecture population to calculate the number of households by family type and by the age and gender of the head of household.

Australia

The Australian household projections are produced using a membership rate method.

Counts of the number of people in each age group in 14 living arrangement types (such as a female parent in a one-parent household or a child in a couple with children) are obtained from the 1991, 1996, 2001 and 2006 censuses. Lines of best fit are used to determine the rate of change in each living arrangement type. The projected living arrangement types are then applied to the projected population to give the number of people in each living arrangement type in each year of the projection period.

The number of families can then be counted from the living arrangement types (e.g. number of one parents families = the sum of the number of male lone parents and the number of female lone parents). The number of families is then converted into the number of family households using the family households to families ratio from the 2006 Census.

The number of non-family households can also be derived from the number of people in certain living arrangement types.

This methodology is used for the national and sub-national household projections with the sub-national projections being constrained to the national projections.

New Zealand

The household projections for New Zealand are produced using a methodology that is very similar to the Australian methodology. There are some differences in the living arrangement types that are used and New Zealand projects the changes in the living arrangement types in two different ways to produce two variant sets of projections; in the first the proportions of the population in each living arrangement type are kept constant at 2006 Census levels and in the second variant a linear projection of the proportions is used.

Annex B – The age/gender breakdown when using different methods

When using a headship methodology, whether using the head of household or household representative person, the following age/gender breakdowns would be available:

for one person and lone adult households with children:
- the number of households by age and gender categories of the adult – cannot tell anything about the age/gender of the children in the household

other households
- the number of households by age/gender of head/HRP - cannot tell anything about the age/gender of the others in the household

Meanwhile if using membership rates are used the following age/gender breakdown would be available:

for one person households
- the number of households by age and gender categories of the adult

for lone adult households with one child
- the number of households by age and gender categories of the adult
- the number of households by age and gender categories of the child

for lone adult households with more than one child
- the number of households by age and gender categories of the adult
- the total number of children of each age/gender category making up the total households
- the average number of children in a household by age/gender category
- the proportion of households with a child of each age/gender category

for all other households:
- the total number of people of each age/gender category making up the total households
- the average number of people in a household by age/gender category
- the proportion of each age/gender category in each household type.