

Small Area Demographic Projections: Methods

Population and household projections for areas within a local authority district

Ludi Simpson, January 2012

This document specifies the methods developed at the University of Manchester to project population and households for small areas within a local authority district of Britain. The methods are also used by Edge Analytics.

1. Data Preparation

1.1. Past population estimates, births, deaths

The relevant data are extracted from source files using a local authority district code. The key datasets available for small area forecasting are as follows:

- Past population estimates for OAs (England and Wales, mid-1991 and mid-2001), LSOAs (England and Wales from mid-2002), and DataZones (Scotland from mid-2001).
- Births by sex and deaths totals for OAs (England and Wales each year from 1991-92), and DataZones (Scotland each year from 2001-02; in Scotland the deaths by age and sex are available).

In England and Wales, LSOA population estimates are disaggregated to OAs using the 2001 Census population. Using the definition of small areas supplied by the user, each OA / DataZone is assigned to a small area. When provided, weights are used to indicate the proportion of each OA/DataZone that falls into each small area.

1.2. Fertility and mortality schedules of age-specific rates

The standard schedules are those from the second year of ONS' latest national projections (for England, Wales or Scotland, whichever contains the small areas; the second year is used, as the first is usually dependent on events recorded in that particular year). These will be used to set the age-sex pattern of rates for each small area. The *level* of fertility and mortality will be specified separately for each small area as described in sections 2.1 and 2.2. In past research, it is the level rather than the age structure of fertility and mortality that most affects the development of births and deaths in local areas.

The ONS expectations of age-sex specific changes in fertility and mortality for the country containing the district are also used, to set the *trend* of future fertility and mortality.

1.3. Migration schedules of age specific rates

The migration schedules described here are required by POPGROUP, but do not have great influence on the final migration patterns used in the projections, which are based on population change in recent years as described in section 2.3.

The balance between migration within the UK and overseas migration is not accurately known for small areas because emigration is not measured even in Census year. Although separate UK and overseas migration streams can be used, they may be misleading as it is the overall balance which is reliably measured.

The 2001 Census provides data for three types of age-specific migration rate for local areas:

- INUK: In-migrants from the rest of the UK
- OUTUK: Out-migrants to the rest of the UK
- INOV: In-migrants from outside the UK1

(There is no OUTOV rate for emigration overseas, since the census does not record those who have moved out of the UK. The schedule of migration rates to overseas (OUTOV) is set to the national age-sex specific rates of immigration from the Census)

Census table ST008 provides rates of migration for each 'Standard ward' (and Postal Sector in Scotland) for 46 age-sex groups. Counts of migrants, based on moves in the year before the Census, are deemed insufficient to provide reliable schedules of migration relevant to each small area. For relatively large areas they may be an alternative source of migration schedules, but for the general approach developed here, four aggregated age groups have been used to represent the main periods of life-time migration.

For age-groups 0-17, 18-29, 30-44 and 45+ the census migration data from ST008 are extracted for each small area: using the ward or wards overlapping the small area with at least half its population (or the one ward with most overlap, if none overlaps with half its population, though this is unlikely except with very small areas). This provides an Observed Migrants file. 'Expected' migrants are

derived by multiplying the small area age-sex population in 2001 by the standard (national) propensity to migrate for each of the four age groups.

Dividing the observed migrants by the expected migrants provides an area factor for each of the flows INUK, OUTUK, INOV for each age category. Multiplying the standard schedule of migration rates at each single year of age and sex by the area factor gives a schedule for each small area that reflects its migration experience in the four age categories. For example, male and female INUK migration rates for each single year from new-born to 17 for a particular area will be multiplied by the same area factor calculated for age 0-17. The resulting schedules are recorded and charted in the migration input files of POPGROUP. The 'Standard' schedule in POPGROUP is computed in the same way for the local authority District, but only used when computing Standardised Migration Ratios in the output for each small area.

2. Training projection – setting assumptions for the main forecast

A first 'training' run of POPGROUP from the base year to the most recent population estimate, uses the births, deaths and population estimates for each small area to estimate recent fertility and mortality differentials, and migrant counts, for each small area. These are input into new POPGROUP files as assumptions for the main projection.

Projections are now typically run with a 2001 base, but where the base year is pre-2001 a constraint to smoothly move each area's total population between 1991 and 2001 is included, rather than a single discontinuity at 2001. Population estimates since 2001 are included as constraints.

2.1. Fertility

The Total Fertility Rate, which in the training run has been calculated by POPGROUP using the birth counts given, is used to calculate differentials. In the main forecast these differentials will be used in the projection after the last year for which births are known.

A weighted average of up to 10 years of Total Fertility Rates for each small area is calculated. To provide a differential for the main projection, it is divided by the standard TFR taken from the most recent GAD projection, which is the standard fertility schedule. The default weights are 0 for years up until the last four years, and a weight of 0.25 for the last four years where births data is available.

The default weights may be changed at the request of the client, if for example one or more of the recent years is thought to be atypical.

Additionally, a further differential is added on the all-groups sheet of the fertility file. This represents the difference between recent years used for small area estimates and the future year used for the schedule.

2.2. Mortality

Similar to fertility, the SMR (all persons) which has been calculated by POPGROUP using the given death counts in the training run is used to compute a differential mortality for each small area. Again these differentials are used in the projection after the last year of data given in the input files.

A weighted average of up to 10 years of the SMR for each of the areas is calculated and then divided by 100 to provide a differential for the main projection. The default weights are 0 for years up until the last four years, and a weight of 0.25 for the last four years where deaths data are available.

The default weights may be changed at the request of the client, if for example one or more of the recent years is thought to be atypical.

Additionally, a further differential is added on the all-groups sheet of the fertility file. This represents the difference between recent years used for small area estimates and the future year used for the schedule.

2.3. Migration Counts – all migration flows

The training projection amends migration each year, at each age and sex, to meet the population estimate at the end of the year. The flows to be amended are specified as 50% each for migration in from the rest of the UK and migration out to the rest of the UK. For example if an initial projection is 10 less than the population estimate for men aged 20, then 5 men aged 20 are added to the migration in-flow and 5 taken from the migration out-flow (*If a migration flow is amended to be less than 0, it is set to zero and the opposite flow increased accordingly. The POPGROUP calculations for implementing constraints in this way are described in the POPGROUP reference manual available at www.ccsr.ac.uk/popgroup/about/manuals.html*).

POPGROUP's 'dump file' records the amended migration. A weighted average of up to 10 years of migrant counts given in the training projection dump file is then used in the main projection for all years after the latest population estimate. The default weights are the same as those for fertility and mortality, for years up until the last four years and a weight of 0.25 for the last four years where population estimates are available. The default weights may be changed, if, for example, one or more of the recent years is thought to be atypical.

The weighted average, calculated for each age-sex category and separately for each of the four migration flows, is repeated for each year of the main projection up to and including the last year.

3. Main population projection: the 'migration-led' projection

The main population projection is called the 'migration-led' projection to distinguish it from the 'dwellings-led' projections where population is affected by housing provision. Such a projection is also more commonly referred to as a 'trend' projection.

The main projection uses the fertility, mortality and migration assumptions estimated as in section 2 above:

- Fertility and mortality differences between small areas will remain
- Fertility and mortality levels will be those of the recent past except for the increases and decreases expected by ONS nationally, and
- The impact of migration for each small area will remain those observed in recent past.

4. Household projection

4.1. Model definition in Derived Forecast model

Small area data used to rescale headship rates in each country usually requires that revised or aggregated household category types are used so that a link can be made between small area census data and household types used in the latest official household projections.

4.2. Estimation of the population not in households

Full information on the population not in households (i.e. communal establishments) by age and sex exists in the census for small areas. These figures are scaled to be consistent with CLG/GRO(S) information for 2001. The counts of population not in households are assumed to remain constant over the projection period. This assumption can be varied on consultation with the user.

4.3. Headship rates for the population in households

'Headship rates' are also known as household representative rates or household representative propensities. Information for small areas exists in the census for the total number of households in each household type, and two broad age groups for one-person households. This is used to scale the official projection district headship rates to local information. Consistency is maintained with the official projection's total number of households in the district in 2001.

4.4. Dwellings/households ratio

Information on dwellings (shared and unshared), households (occupied, second homes and vacant) are collected for each small area from the 2001 Census. A single household/dwelling ratio is used on the DFSupply input workbook, to allow conversion between households and dwellings.

4.5. Housing-led projections

Each (optional) scenario of new development that has been provided by the user is entered on a separate constraint (CONS) workbook, in years after the population constraints. Migration flows are amended so that the population in each projected year is consistent with the number of households implied by the development scenario (50% of the necessary amendment is made to each of the two UK migration flows).