

2006-based Subnational Population Projections for England

Methodology Guide

1. INTRODUCTION

This short guide describes the methodology used in producing the subnational population projections for England.

Subnational population projections are produced to provide estimates of the future population for the English regions, local government and strategic health authority areas based on common underlying assumptions. The projections are also produced for Primary Care Organisation (PCO aka PCT) areas. The projections are produced every two years by the Office for National Statistics (ONS) to coincide with the production of the National Population Projections (NPPs). The most recent national population projections, published 23 October 2007, are 2006-based projections and use recent past demographic data to form assumptions about future demographic behaviour.

The projections form a consistent basis as an input to household projections and for use in local planning of health and transport service provision. They are also used by local authorities (together with other factors) to produce their own projections for local planning and recently by the Department of Health and Department for Communities and Local Government (CLG)¹ for funding allocation to PCOs and local authorities. At the England level they are fully consistent with the 2006-based population projections for England that were published by ONS on 23 October 2007².

Overview of Methodology

The projections take as their starting point the 2006 mid-year population estimates published on 22 August 2006³. The projections for each year are calculated by ageing on the population from the previous year, applying local fertility and mortality rates to calculate the projected number of births and deaths, then adjusting for local area migration into and out of the area. This process is repeated for each of the twenty-five years of the projection period. The sum over all areas of the projected population and births, deaths and migration components are constrained to the totals for England in the 2006-based national population projections produced by ONS.

Projections for PCOs are based on those for Local Authorities. In many cases PCO areas are coterminous with local authorities or aggregations of local authorities, in which case projections for these areas are calculated by aggregating the appropriate local authority level projections. Where areas are not coterminous, PCO projections are produced by apportioning local

¹ 2004 based projections are currently used by CLG for funding allocation to local authorities. It is not currently planned to use 2006 based projections for this purpose because of the three year funding settlement cycle.

² Historically produced by GAD, but from January 2006, the National Population Projections are produced by the Office for National Statistics Centre for Demography.

³ Available at <http://www.statistics.gov.uk/popest>

authority level projections, based on 2006 estimates of PCO populations. This is done by single year of age and sex.

2. DATA SOURCES

The following data are used to calculate subnational population projections by single year of age and sex for all areas:

- Small area (building brick) data for base population, births, deaths, internal and international migration (in, out and net migration flows);
- National population projections assumptions for England; population, births, deaths and international migration flows

The small area (building brick) level data conforms to an ONS standard set of geographic units, which correspond to Local Authority districts.

2.1. Base Population

The projection calculation starts with a base population. The data used are at single year of age and sex and are available at the building brick level. The population is divided into three types for the purposes of projection:

- the general population
- home armed forces
- foreign armed forces and their dependants

2.1.1. General population

The ONS 2006 mid year population estimates for England provide the base data for the general population. These are used at single year of age and sex level and are constructed by rolling forward the previous years mid-year estimate and adding in changes for births, deaths and migration and any other changes. The estimates and their methodology are fully described in the guide "Making a population estimate in England and Wales"⁴.

The general population covers the normal resident population at their usual place of residence. This includes all those temporarily away from home (for six months or less) and excludes visitors. Home armed forces stationed outside England are not included. Asylum seekers and visitor switchers (people who enter the country intending to visit, but stay to become usual residents) now residing in England are included as part of the general population. Students are included as taken to be resident at their term-time address.

2.1.2. Home and foreign armed forces

Home and foreign armed forces constitute two separate population types and are treated as separate static populations in the projections model. Data on UK armed forces originate from the Defence Analytical Services Agency (DASA) and data on foreign armed forces originate mainly from US Air Force

⁴ Available from: <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=575>

statistics. They include numbers of home and foreign armed forces and their dependants living in barracks and in army quarters.

2.2. Local level births and deaths

Registered births for 2002-2006 by age of mother are used to create age-specific fertility rates (ASFRs) for each sex, building brick and year. These are compared to the national fertility rates used in the national population projections and adjusted to create local age specific fertility rates for each year of the projection period.

Deaths are treated in a similar way. Registered deaths for 2002-2006 by single year of age and sex are used to create age-specific mortality rates (ASMRs) for each sex, area and year of data, which are then averaged and compared to the national mortality rates and adjusted to create local age-specific mortality rates for the projection.

2.3. Migration

These data are the most complex and require careful restructuring for use in the projection calculations. This partly reflects the limitations of the data sources available. While there is a legal requirement to register a birth or a death there is no such requirement to register a change of address. The subnational projections model uses migration data for single year of age and sex for each building brick as with the other components. Estimates are required for in and out migration flows for both internal and international migration.

2.3.1. Internal Migration

For subnational population projection purposes this is defined as migration within England only⁵. From 1999 the key source of information on internal moves has been from the General Practitioner (GP) patient records. Every health authority in England holds a register of the patients registered with the GPs within their area of responsibility. ONS receives an annual extract of this information from which it creates a total patient register for the whole country. Comparing successive registers allows estimation of moves within the country. An internal migrant is defined as someone who changes their area of residence between one year and the next.

This method of estimation will miss certain groups of people who do not appear on the registers in two successive years (births, deaths, those joining or leaving the armed forces or entering or leaving the UK) and so the number of migrants is constrained to the National Health Service Central Register (NHSCR), which measures moves between former health authorities only

⁵ For other uses internal migration will include flows to and from Wales, Scotland and Northern Ireland, termed cross-border flows. However, the propensity to migrate model used to project internal migration requires a full matrix of flows out of and into each LA by single year of age and sex. This level of detail is not required to project cross border flows and therefore projections add these flows to other international migration flows.

(rather than the postcode level information) but has the benefit of being constantly updated.

The internal migration data provides a large matrix of moves by origin area, destination area, sex and single year of age, which is subsequently used to create age profiles of migration and calculate the propensity to migrate for each area. The internal migration flows within the matrix sum to zero.

2.3.2. International Migration

The projections model treats international migration as being made up of three separate types of movement:

- Between England and the rest of the UK (deemed part of internal migration for UK-wide statistics)
- Between England and the rest of the world
- Visitor switchers and asylum seekers

The data sources for these are summarised below. Local estimates are calculated from these data by disaggregation (using supplementary information including data from the Labour Force Survey and Census) and these are then used to distribute the national assumptions on international migration as used in the national population projections to the smaller area level.

Migration between England and the rest of the UK is calculated in the same way as the internal migration assumptions. Five years of the GP patient register data giving flows between English building bricks and areas of Scotland, Northern Ireland and Wales are used to infer levels for the projection period.

The main source of information on international migration data for the rest-of-the-world stream is the International Passenger Survey (IPS). This is a sample survey of passengers travelling through airports, seaports and the Channel tunnel. It provides information on the number of people intending to stay in the UK or leave the UK for a year or more, at a national and former health authority level. Since the estimates of international migration from the IPS are subject to some uncertainty (as measured by the sampling error) the last six years' data are combined. The way these data are combined has been change for 2006-based subnational population projections due to the improved methodology used by mid-year estimates. Further detail on how these data are averaged is given in Annex 1: Calculation of an Average for the International Migration Component of Change.

Visitor and migrant data from the IPS are used to estimate changes in intentions. Visitor switchers are people who initially visit or leave the UK for a short period but subsequently stay for a year or longer and therefore become migrants. In contrast migrant switchers are those who intended to migrate but in fact stay in the UK or abroad for less than one year and should not be counted as migrants.

Data on asylum seekers and their dependants are provided by the Home Office and the National Asylum Support Service. Applications for asylum (excluding an estimate of those removed from the UK within one year and a small number of asylum seekers captured by the IPS) provide the basis for estimated inflows of asylum seekers. Data on removals, refusals, withdrawals and appeals for principal applicants and dependants are used to estimate outflows of unsuccessful asylum seekers leaving the UK without being captured by the IPS.

Estimates of migration flows between the UK and the Irish Republic are produced using information from the Irish Quarterly National Household Survey and the National Health Service Central Register.

The ONS recognises that there are limitations in the data and methodology used to distribute international migrants within the subnational population projections. Improvements in these distributions have been researched as part of the ONSCD 'Improving Migration and Population Statistics (IMPS)' project⁶. The first improvement outcomes of this work programme were incorporated into the 2006 mid-year population estimates (published on 22 August 2007⁷), including revisions for the back series of mid year estimates for years 2002 to 2005. The 2006-based subnational population projections take full account of these improvements.

2.4. National Projection Data

The subnational population projections are fully consistent with the 2006-based population projections that were produced by ONSCD in consultation with the Registrars General. The national projections for England were published by ONS on 23 October 2007⁸.

The national level data on births, deaths and international migration for each sex and year of age are required as a control. Local data are therefore the main basis for the projections for each area but each component and the projected population is scaled to ensure that the aggregated data equal the national population, birth, death and migration projections. At local level this means that figures may be scaled up or down.

3. PROJECTION METHOD

In brief the projections use a recognised cohort component methodology. They take as a starting point the population in 2006 for each area by single year of age and sex. This population is then projected one year ahead by:

- Ageing on the population

⁶ Further information on the IMPS project can be found at:
<http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/default.asp>

⁷ <http://www.statistics.gov.uk/popest>

⁸ See: <http://www.statistics.gov.uk/CCI/nugget.asp?ID=1352&Pos=6&ColRank=2&Rank=448>

- Applying age-specific fertility⁹ rates to derive a number of births
- Applying age and gender specific mortality rates to derive a number of deaths
- Adjusting for the expected number of people entering and leaving an area by age and sex
- Controlling the sum over all areas to the national population projections

This process is then repeated for the 25 years that the projection covers.

The basic equation for the projection can be written as:

$$P_p = P_c + B * S_1 + B * S_2 - D + M_w + M_i$$

Where

P_p = projected population

P_c = current population

B = projected births

S_1 = sex ratio of males

S_2 = sex ratio of females

D = projected deaths

M_w = net internal migration from within England

M_i = net international migration from outside England

Separate population types - foreign and home armed forces - are dealt with separately from the general population and assumed to remain static for the projection period.

4. STAGES IN THE MODEL

4.1. Ageing on

The population in each area is aged on one year to become the appropriate age in the following year of the projection. For example 17 year olds in 2006 will become the 18 year olds for 2007.

4.2. Adding Births

Local birth differentials are calculated by using the latest five years of birth data by age of mother to create local age specific fertility rates. The projected number of births is then calculated by multiplying these age specific fertility rates by the number of women in each age group in the area.

The projected total births are then split by sex using a sex ratio at birth derived from the national totals of historic births data.

Finally the total number of births is controlled to the national projected total for births by applying the appropriate scalar derived from a comparison of the

⁹ That is the age of the mother.

national projection birth total with the sum of the births over all areas from the projected figures.

4.3. Subtracting Deaths

A similar method is used to calculate the estimated number of deaths for each age and each projection year. Local death differentials are used to calculate death rates by age and sex. These are then applied to the population of an area to produce a projected number of deaths by age and sex, and the total across all areas is constrained to the national projections in the same way as the births.

4.4. Net Internal Migration Adjustments

Information for modelling the internal migration component of the methodology is derived from the GP registers (PRDS) source. This holds postcode level information and so provides more local information on origins and destinations than the central register information (NHSCR) as used in earlier sets of projections.

From this we build a large multidimensional matrix of recent moves by origin area, destination area, sex and single year of age for the five historic reference period years on which the assumptions about future levels of migration are based.

The modelling and calculation of internal migration follows a set of processes:

i) Assignment matrices are calculated. These show the proportion of total migration to all other areas of England for every building brick in the projection set. In a small number of areas these migration rates are implausibly high, consequently single year of age propensity to migrate rates are capped at 0.8 for females and 0.6 for males.¹⁰

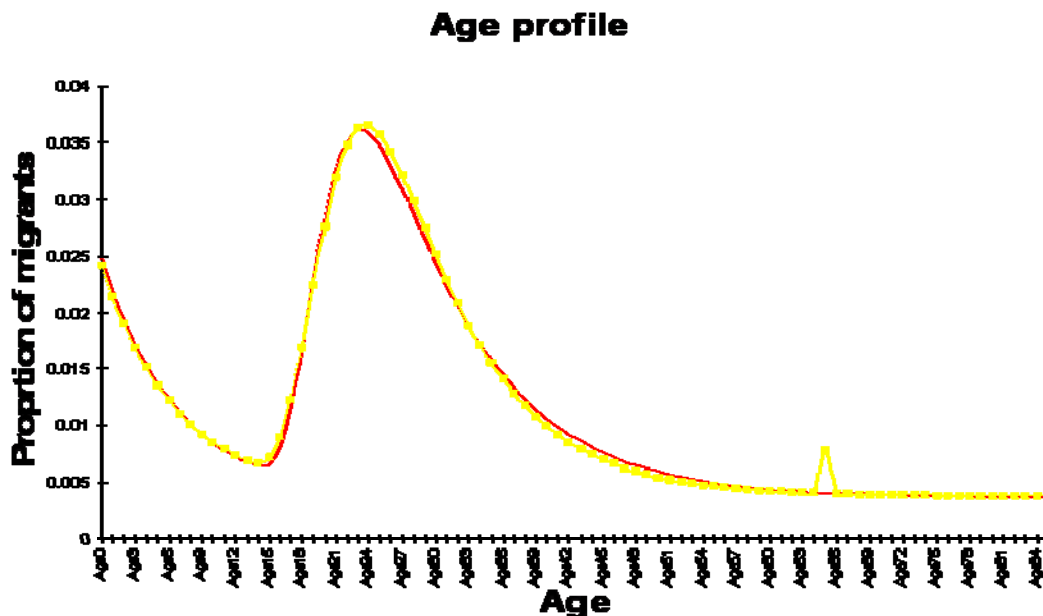
ii) The out-migration likelihood by age and sex is calculated by fitting a family of curves to the historic data. The curves represent a known and researched pattern typically found in the age profile of migrants from a population. Commonly known as the Rogers curve and based on the work of the International Institute for Applied Systems Analysis (IIASA) (Castro and Rogers, 1979), it represents moves that can be broadly categorised into three groups most likely to migrate:

- Young adults aged between 17 and 28 who are likely to migrate either leaving the parental home, moving to study or as part of a highly mobile component of the labour force;

¹⁰ Further details on why these rates are capped and which areas are directly affected in 2006 based projections can be found in: *2006 based Subnational Population Projections: Summary of Part One Consultation Response Analysis and Changes Implemented in the Projections*. This document can be found from here: www.statistics.gov.uk/snpp

- Family migrants including parents with children, usually aged between 25 and 59 with offspring aged up to 16 years;
- In some areas there will also be a small hump in the profile around retirement age to reflect those who move at or near retirement.

For example:



This curve can be expressed in terms of six parameters relating to the ascending and descending parts of the peaks of maximum migration for these groups, and to the general level of migration of the area. In most cases the age profile shape remains relatively standard but the general level of migration varies as measured by the Gross Migration Rate (GMR).

The age profile curves are fitted for every area and for both sexes separately.

Because of distortion by small number effects migration curves for the Isles of Scilly and Rutland are replaced by those for, respectively, Carrick and Winchester.

iii) The out-migration probability as modelled in the previous step is applied to the current estimate of the population in each area. This generates an assumed number of out-migrants for the projected year by age and sex.

iv) The destination of these out-migrants is calculated. This is obtained by applying the assignment matrices of origins and destinations as calculated in step 1 to the total out-migrants for the area. For any area we are able to allocate proportions of the total assumed out-migration to every other area as appropriate.

Totalling this up for each destination area in turn then provides an assumed inflow by age and sex, which when totalled across all areas in England equals the total assumed outflow.

v) The net migration adjustment for each area is calculated as the difference between the assumed inflow and the assumed outflow. The total net internal migration adjustment across all areas of England naturally sums to zero.

4.5. Net International Migration Adjustments

The key assumptions on the levels of international migration by age and sex are taken from the assumptions on these used in the official National Population Projections. The streams of international migration are:

- International migration to and from the rest of the UK (“cross-border flows”)
- International migration to and from the rest of the world
- Asylum seekers
- Visitor switchers
- Irish flows

Assumptions on the national levels of these are calculated using historic data sourced from the International Passenger Survey (IPS) and data from the Home Office and National Asylum Support Service (on asylum seekers).

Each stream of international migration is disaggregated to the subnational level on the basis of building-brick level historical data.

For the Rest of UK stream the PRDS data is used in a very similar way as for the internal migration calculations – the PRDS gives us information on the origins and destinations of moves including those from Scotland, Northern Ireland and Wales and using these records it is possible to proportionally allocate the GAD assumptions on cross border flows to each area of England.

For the Rest of the World stream IPS area data is further disaggregated to the building brick level using the new method as used in the latest mid year estimates. Inflows are distributed based on an updated method that takes into account a new intermediate geography and makes use of Labour Force Survey data¹¹. Outflows are based on a ‘propensity to migrate’ model that takes account of demographic features of the existing population in each area¹². Note that this method differs from that used in sets of projections prior to 2004-based revised projections, where 2001 Census information on moves with the origin of the move outside the UK was applied directly to the IPS aggregate results. Six years of data are used for rest of the world migration, the reasons for this and further details of how these data are averaged are

¹¹ http://www.statistics.gov.uk/downloads/theme_population/LFS.pdf

¹² http://www.statistics.gov.uk/downloads/theme_population/PtM2002_6.pdf

given in Annex 1: Calculation of an Average for the International Migration Component of Change.

The visitor switcher and asylum seeker assumptions are disaggregated as per the method used for compilation of the Mid Year Estimates¹³. For both of these streams one year of data are used. For asylum seekers this is because the flows for this component have changed rapidly in recent years. For visitor switchers this reflects that data on this component for the new methodology have only been collected in the IPS for a limited period of time.

4.6. Additional Migration Redistributions

The penultimate stage in the projections process is to make any necessary adjustments to the migration assumptions at the local level.

Since the migration element is the hardest component of change to measure the ONS carried out a consultation exercise with Local Authorities, PCOs, Strategic Health Authorities and regional bodies between 6 March 2008 and 17 April 2008. Authorities were supplied with the proposed migration elements and asked to comment on whether there had been any local issues during the reference period that might have resulted in an atypical level of migration for that time, and subsequently impact upon the assumed trend for the projected period.¹⁴

4.7. Final Population Control

Following the calculation of the various stages in the model each component is summed and controlled to the national totals (from the ONS national projections). Additionally the final projected population is controlled so that the sum across all areas matches the National Projection population at single year of age and sex.

This step-by-step process results in a projected population for the first year of the twenty five years processed. This first year projection then forms the base population for the next year's calculation and the process is repeated twenty five times.

Projections Branch
Office for National Statistics Centre for Demography
June 2008

¹³ See section 2.6.4 of "Making a Population Estimate in England and Wales", available from <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=575>

¹⁴ Further details on changes resulting from the consultation can be found in *2006 based Subnational Population Projections: Summary of Part One Consultation Response Analysis and Changes Implemented in the Projections* available from here: www.statistics.gov.uk/snpp For 2006 based projections specific adjustments in this context were made only for Forest Heath.

Annex 1: Calculation of an Average for the International Migration Component of Change.

Background

In the revised 2004-based projections an average of the past five years had been used as international migration data straddled the change in the mid-year population estimates methodology for estimating international migration. However, now that all five years of data are on the new methodology (which makes use of three year averages) we had decided to use an average of two 'three year averages'. Previous research shows that to use an average of the estimates for the past five years, when each estimate is itself a three year average, weights centre years too much and underweights end years. Of particular importance is that this underweights the most recent year.

This annex deals with the IPS component of international-migration only. Visitor switchers, asylum seekers, flows to and from the Republic of Ireland, and flows to and from Wales, Scotland, and Northern Ireland are dealt with separately and are not subject to the same concerns.

Calculation Levels and Averaging Periods for Annual Mid-Year Estimate Calculations:

The table overleaf gives a summary of the levels involved in calculating the migration estimates for mid-year estimates that then feed into subnational population projections.

IPS = International Passenger Survey
 GOR = Government Office Region
 LFS = Labour Force Survey

	In	Out
UK	1 year IPS	1 Year IPS
England	As for GOR – 3 year LFS average	1 Year IPS
GOR	3 year LFS average, y-2, y-1, & y, Spring Quarter only, technically as this is based on a calibration method this is an unbiased estimator of a three year average – but any differences will be very small.	1 year IPS
Intermediate geography	Outside London: 3 year IPS average London: Students Census, non-Students 3 year average LFS data (similar to GOR)	3 year IPS average
LA	Census	Single year model (including fixed Census and variable updated data of which data on in-migrants is especially important)

Considerations:

Ideally we would operate the models for in- and out-migration at a five year average period from first instances for population projections. However, the current implementation of the models for estimating the international migration component of change does not facilitate this. Amending the supporting systems to allow direct five year averaging is a major project that will take considerable time and resource. However, we do note this should be an area for development prior to the next projections round.

To avoid creating a weighted average effect in the numbers we need to use proportions for averaging rather than absolute numbers.

As much of the process will rely on two sets of three year averages (providing a six year average), where possible we will use six year averages throughout. This avoids unbalancing in- and out-migrants.

The absolute numbers in the input data do not ultimately matter for the projections model as total international migration flows are constrained to National Population Projection assumptions for England. Therefore, it is the proportionate allocation that is important; nevertheless proportions will be applied to the average of 2001 to 2006 national figures to provide numbers that facilitate ease of comparison.

Methodology

In-Migrants

- a) Simple average of 2003 and 2006 GOR proportions of UK figure.
- b) Simple average of 2003 and 2006 LA proportions of GOR (use of Census at final level mean that consideration at intermediate geography is not needed).
- c) Apply proportions at a) to UK 2001 to 2006 average figure to give GOR figures.
- d) Apply proportions at b) to GOR totals at c to give LA totals.

Out-Migrants

- a) Simple average of GOR proportions of England figure for 2001, 02, 03, 04, 05, and 06.
- b) Simple average of intermediate geography proportion of GOR for 2003 and 2006.
- c) Simple average of LA as proportion of intermediate geography for 2002, 2003, 2004, 2005, and 2006 (note no model available for 2001).
- d) Apply GOR proportions at a) to 2001 to 2006 Average England Figure.
- e) Apply intermediate geography proportions at b) to GOR totals at d)
- f) Apply LA proportions at c) to intermediate geography totals at e) to give LA figures.

Age and Sex

LA total figures are calculated 'independently' from age/sex distributions. Therefore a simple averaging of proportions can be applied to figures by age and sex. Separately for ins and outs calculate the simple average of each age/sex cell as a proportion of the LA total, e.g. M0, over 2002, 03, 04, 05, and 06 (again no model available for 01).

Conclusion

These processes provide the closest possible approximation to a six-year average – and take into account the multi-stage model process shown in the table above.