

2012-based Subnational Population Projections for England

Report on Unattributable Population Change

20 January 2014

Introduction

The Subnational Population Projections (SNPPs) are produced using assumed future trends in the components of population change. These trends are calculated using the best series of data from the most recent 5 or 6 years up to the base year of the projection. The revised back series of population estimates from mid-2002 to mid-2011 contains an additional component of population change known as Unattributable Population Change (UPC). ONS recognises that the UPC component can have a significant impact at some age/sex groups in some local authorities. Work has taken place to investigate the impact of UPC (Appendices A & B) and how an adjustment may be made for it (Appendix C).

It is proposed that no adjustment be made in the 2012-based Subnational Population Projections for the unexplained component of population change in the revised population estimates series. This paper describes what UPC is and why we are not making an adjustment for it in the 2012-based Subnational Population Projections.

What is Unattributable Population Change?

Following the 2011 Census, the intercensal population estimates were rebased so that the mid-year estimates (MYEs) for the period mid-2002 to mid-2011 were in line with the 2011 Census. After making allowances for methodological changes and estimated errors in the components during the decade, the remaining difference between the rolled forward 2011 MYEs and the 2011 Census based MYEs for England was 103,700. This is referred to as Unattributable Population Change (UPC) in this paper. In order to produce the revised series of population estimates for the last decade, the UPC was apportioned across each of the 10 years using the cohort method which takes account of the fact that individuals age as the decade progresses. This method was applied to both the national and subnational MYEs. An example of how this method works can be found in the paper on the methods used to revise the subnational MYEs at www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru/latest-publications-from-the-population-statistics-research-unit/index.html .

Some of the difference between the rolled forward 2011 MYEs and the 2011 Census based MYEs for England has been explained by the following:

- EU8 immigration adjustment
- Republic of Ireland migration revisions
- Migrant switcher revisions
- Visitor switcher revisions
- Armed forces adjustment
- Cross-border migration correction
- Mid-2009 asylum seekers and visitor switchers correction
- Removal of historic processing adjustments

The migration data used in the 2012 SNPPs include the adjustments described above that were made to the population estimates revised after the 2011 Census.

Appendix A shows how the UPC is distributed both nationally and by local authority. At the national level, the UPC of 103,700 affects some age groups more than others. At the subnational level, the UPC affects some local authorities (LAs) more than others. However, the age distribution of this difference at national level is not always the same as that observed subnationally.

Potential sources of Unattributable Population Change

The UPC is likely to be due to a combination of sampling variability, or other issues, in the following:

- International migration estimates
- Census estimates, both 2001 and 2011
- Internal migration estimates (at subnational level only)

Further detail on how these components may have caused the UPC can be found in Appendix B.

How UPC was handled after the 2001 Census

The final unattributable population change after the 2001 Census (including adjustments made to mid year estimates to correct for census deficiencies) was 209,000. Initially the figure was higher and a UPC adjustment in the mid year estimates was made to the mid-2002 estimates. The final level of 209,000 (just 20,000 a year) was considered too small to make an adjustment for in the MYEs for the following decade. This UPC of 209,000 for England and Wales was double the UPC of 103,700 for England after the 2011 Census.

Justification for the decision not to make an adjustment for UPC in the 2012-based Subnational Population Projections

No adjustment is to be made in the 2012-based Subnational Population Projections for the unexplained component of population change in the revised population estimates series. An adjustment for UPC could only be made if it can be demonstrated that it measures a bias in the trend data that will continue into the future.

Quality assurance of the 2012-based Subnational Population Projections did not reveal any problems indicating that adjustments for UPC are necessary. The resulting projections generally appear to better reflect trends across all the LAs than recent sets of projections.

ONS decided not to make an adjustment for UPC in the 2012-based National Population Projections or in the series of population estimates based on the 2011 Census. This is because the UPC for England (103,700) is within the confidence interval for the international migration estimates. It is also within the sum of the confidence intervals for the 2001 and 2011 Census.

The UPC is unlikely to be seen in continuing subnational trends as:

- It is unclear what proportion of the UPC is due to sampling error in the 2001 Census, adjustments made to MYEs post the 2001 Census, sampling error in the 2011 Census and/or error in the intercensal components (mainly migration).
- If it is due to either 2001 Census or 2011 Census then the components of population change will be unaffected
- If it is due to international migration, it is likely that the biggest impacts will be seen earlier in the decade and will have less of an impact in the later years, because of improvements introduced to migration estimates in the majority of these years.

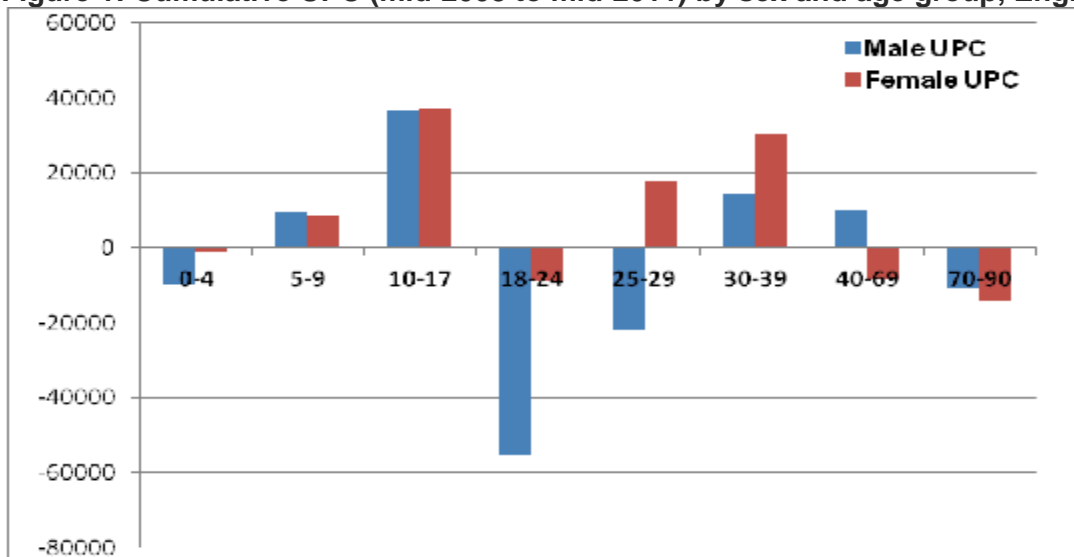
Therefore ONS propose that no adjustment be made in the 2012-based Subnational Population Projections for the unexplained component of population change in the revised population estimates series.

Appendix A - How the UPC is distributed

At national level

At national level (England) UPC can result from uncertainties in the population numbers at 2001 and 2011, as well as issues with estimation of cross border and international migration. Figure 1 shows how the cumulative UPC for England is distributed by age and sex. For ages 10 to 17, the positive UPC suggests an underestimation in the rolled-forward mid-year estimates (MYEs). This may be due to a possible underestimation in the 0 to 9 age group in the 2001 Census and/or an underestimate of international migration. Young men aged 18 to 29 show a negative UPC which means that there was an over estimation of young men in the rolled-forward MYEs. Those aged 30 to 39 have a positive UPC and so appear to be underestimated in the MYEs. The elderly age groups show a small negative UPC which suggests an overestimation in the MYEs.

Figure 1: Cumulative UPC (mid-2008 to mid-2011) by sex and age group, England



At subnational level

To illustrate how the average UPC is distributed across LAs for a selected sex and age group we can take look at an example. Figure 2 looks at males aged 20-24 and illustrates that the national pattern is not observed across all areas subnationally. It can be seen that there are LAs at both ends of the graph that have a larger UPC. This pattern can be seen for most age groups at the local authority level, with some LAs showing a positive UPC and some LAs showing a negative UPC.

However, the subnational distribution is not consistent with the distribution of UPC at the national level. For example, at the national level, the UPC shows that the 10-17 age group was underestimated in the MYEs and that the 18-24 age group was over estimated in the MYEs. However, at a local level, this under/overestimation does not always apply. There are many LAs where all age groups appear to be either over or underestimated in the MYEs. For example, in Leeds the UPC for males indicates an overestimate in the MYEs across all age groups and in Liverpool, the UPC for males suggests an underestimate in the MYEs across all age groups. This is illustrated by figure 3 below.

Figure 2: Average UPC (mid-2008 to mid-2011) for males aged 20-24

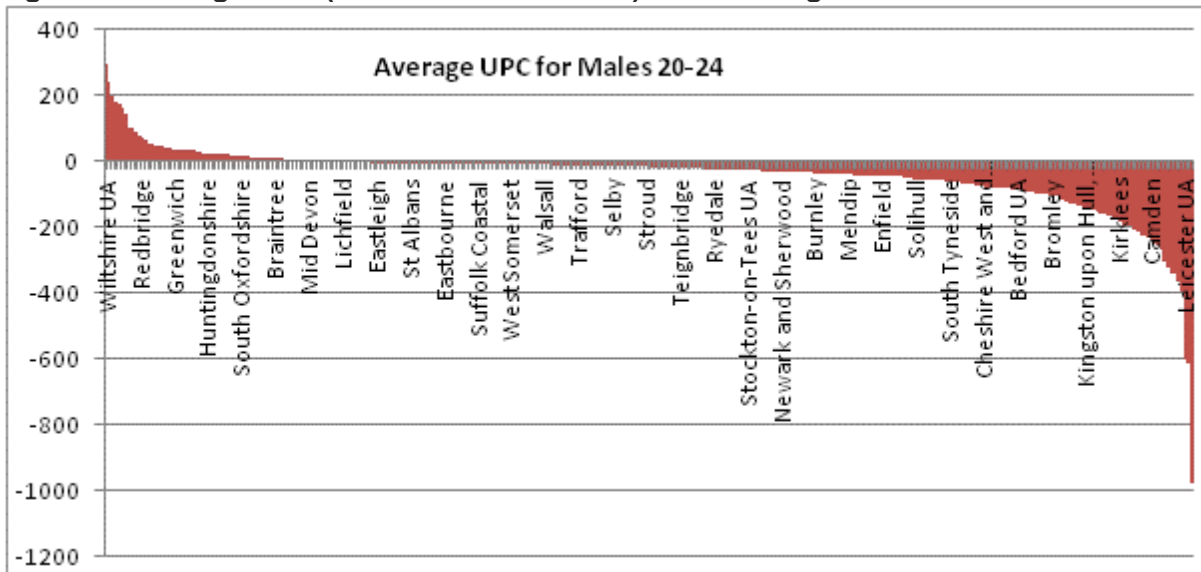
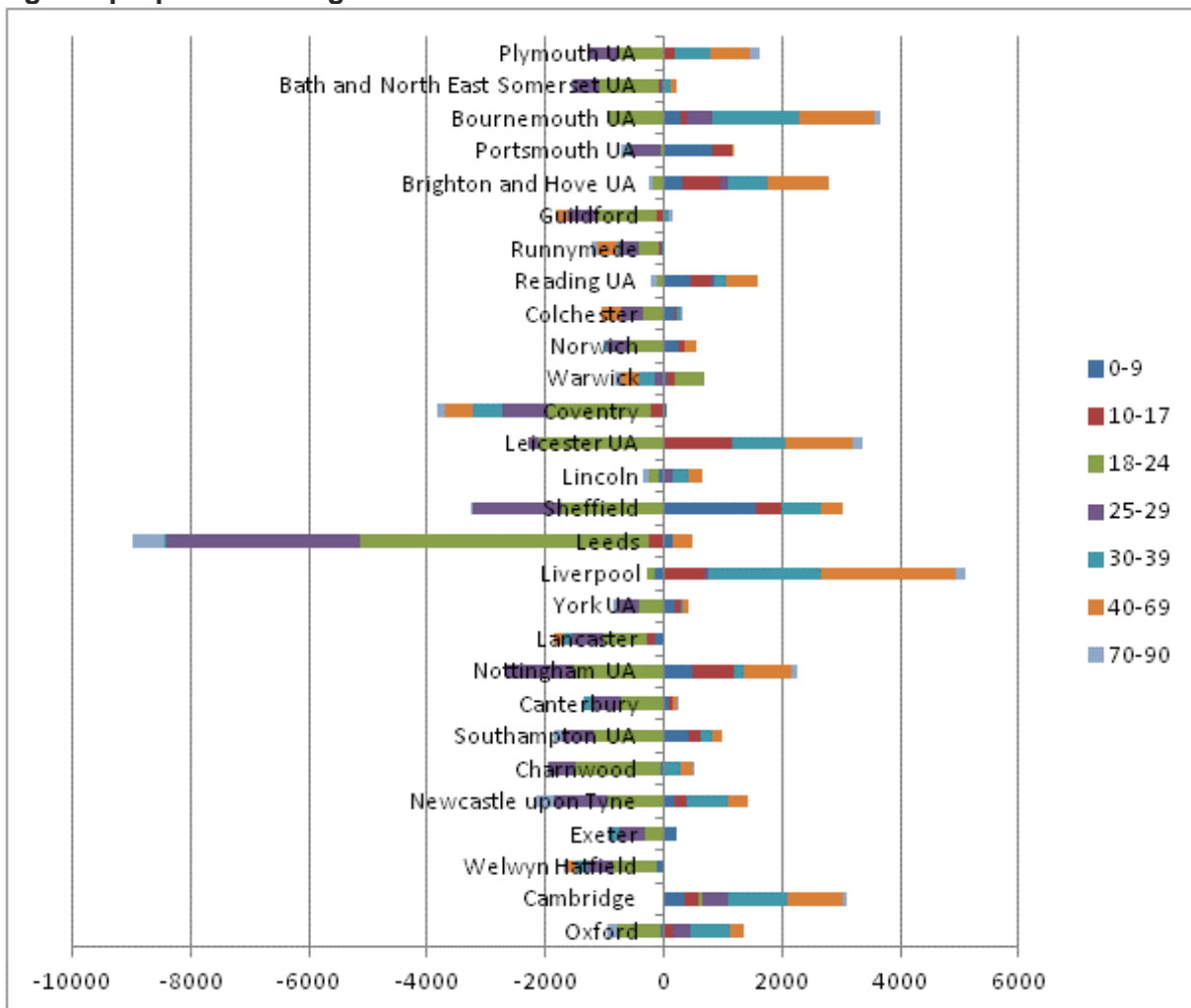


Figure 3: UPC for males (mid-2008 to mid-2011) by age group for LAs in England with the highest proportion of Higher Education Students



Appendix B - Possible causes for UPC

The UPC is likely to be due to a combination of reasons. These possible causes are summarised below:

2001 Census

The 2001 Census is the base for the population estimates over the decade. Although adjustments were made to these to improve the estimates, errors may still remain in the base estimates either due to sampling error from the census estimation process or from the adjustments made to improve the population estimates base following the 2001 Census.

2011 Census

The 2011 Census which is used as the benchmark for the mid-11 population estimates is subject to sampling error. Although the 2011 Census estimates are considered to be of high quality, the need to estimate the number of people who did not appear on a Census form means that each LA inevitably has some uncertainty around their estimate. There is also the potential for other unidentified biases to have occurred. However, any error in the 2001 or 2011 Census that is contributing to the UPC would not continue into the next decade. This is because the error would not be part of the components of change which are used in the trends for projecting the population.

International migration

Both the immigration and emigration estimates are subject to sampling errors. In addition, the new methods for distributing immigration down to LA level have only been applied to the years ending mid-2006 onwards. If it had been possible to apply the new methods to the earlier years in the decade, this would have led to different estimates for each LA. In addition to this uncertainty, applying an age and sex distribution to international migration estimates at the national level is subject to error.

Internal migration inaccuracies

Apart from changes to school boarder estimation and small revisions to cross-border flows, migration between LAs in the UK has not been changed in the back series. However, there is likely to be some level of inaccuracy in the internal migration estimates over the last decade as some moves are difficult to estimate accurately. For example, we know that the movement of young people finishing Further Education courses is difficult to capture. Previous research has found long time lags between moving and re-registering with a GP in the student age groups. This has resulted in moves being captured in the wrong time period but also at the wrong age, so there will be too few moves at younger ages (the age of moving) and too many moves at older ages (the age of re-registering). It is not known whether these known lags in internal re-registrations may change in the future. A further issue related to lagged moves is that in the years following the Census, a small minority of moves may be estimated which have already been accounted for by the taking of the Census.

Definition of a prisoner

The definition of a prisoner has changed. The mid-year estimates up to mid-2010 define a prisoner as someone who has already served at least six months in prison by the mid-year point. However, the mid-2011 estimates have moved to a new definition that a prisoner is someone who is on a

sentence of six months or more, regardless of when their sentence commenced. The overall impact of this is that more people will be defined as prisoners, which will increase the population of LAs with prisons, potentially by several hundred and slightly reduce it in other LAs. This will have a small effect in those LAs with a large prisoner population.

Appendix C – Research carried out to identify whether it would be possible to create an adjustment for UPC in the SNPPs

ONS are aware that some local authorities believed that UPC would be an issue in calculating trends for the SNPPs. Therefore, before data were available to produce the projections, ONS considered how an adjustment could be made within the limitations of the SNPP processing system. A possible solution was identified. However, more research would be required on the specific details of a method to put into production.

It was decided that the method used should not make an adjustment to any of the current migration components as this would not be transparent to users and would require evidence that the UPC was due to a particular component which is something that is not known.

Methodological Constraints

Any method to make adjustments for UPC in the 2012 SNPPs has to work within the constraints of the SNPP processing system.

Firstly, any adjustment would have to be processed through the now redundant visitor switcher flows in the system. The impact of this is that any adjustment made will have to remain constant over the whole projection period. The advantage of using this flow is that it can be separated out for creating outputs if required.

Secondly, as the SNPPs are constrained to the national population projections; and there is no UPC component in the national population projection, any adjustments made to the SNPPs will have to sum to zero across all local areas for every age and sex group. The impact of this is that any adjustment for UPC at the subnational level will just move population around from one LA to another. This means that if we add people to one LA, we have to take them out of another. As the UPC does not sum to zero across all local authorities in England, this would mean that any adjustment could potentially be biased.

Therefore the method to create adjustments by age and sex for each local authority must satisfy the constraint that the adjustments sum to zero, and will be constant for each year of the projection. A general approach has been considered on how these adjustments could be created. However, the finer detail on exactly how these methods would be used in practice has not, and would require further research.

Methodology

The general approach would be to create the adjustments using the UPC component from the revised population estimates, for the years from mid-2007 to mid-2011. The method scales the average UPC by age and sex for some LAs so that across all LAs, the UPC adjustment would sum to zero. This approach may be best applied to five-year age groups and then the average scaled UPC for each age group can be apportioned down to single year of age

As UPC by age does not sum to zero, in some areas all of the average UPC will be adjusted for and in others this will be only a proportion. Figure 2 in Appendix A shows the UPC for males aged 20 to 24 years old. It is clear the negative UPC is much greater than the positive UPC. The method would scale down the average in the areas where the UPC is negative, but the average in areas where the UPC is positive would be unchanged.

Even if it were accepted that there was an element of UPC in some age groups and in some LAs that would continue into the future it would not apply to all LAs; many LAs and some age groups are not impacted greatly by UPC. An alternative approach would be to restrict any adjustments to a subset of age groups or areas. However the overall approach would remain the same. It would be straightforward to adjust the method for a subset of age groups.

If the adjustment was to be made on a subset of areas then a criteria on selecting these areas would need to be set so that the method is applied consistently across LAs in England. The criteria would likely be set based on percentage of UPC in an area for a given age group. However it is very possible that the criteria only selects LAs with a negative UPC, in which case as the UPC has to sum to zero no adjustment would be possible.