

Population And Migration Statistics (PAMS) Committee (Scotland)

Small Area Population Estimates Review

1. Introduction

1.1 This paper provides a summary of the work carried out looking into alternative ways of producing the annual small area population estimates (SAPE). PAMS members views are welcome.

2. Recommendation

2.1 After investigating various different methods for producing the SAPE we recommend that the [cohort component method](#) as currently used, is also used for the small area population estimates, based on 2011 data zone boundaries..

3. Different options

3.1 The small area population estimates rely heavily on the mid-year population estimates methodology for council areas and any alteration to the SAPE methodology would have to be carried out prior to the Scotland, Council area and NHS Board mid-year population estimates publication.

4. We looked at four different methods.

4.1 **Cohort Component Method** – the cohort component method involves the ageing-on the population age structure from one year to the next, then applying counts of births and deaths that take place during the year and adding on estimates of migration for each small area. Special population groups (armed forces and prisoners) are removed from the population at the start of process and updated counts of these groups are added towards the end of the process. This method is currently used to produce the Council area and small area mid-year estimates in Scotland.

4.2 **Ratio Change Method** – The ratio change method uses an indicator of the current and the past population to update an earlier population estimate for each small area. The assumption is that the relationship between the indicator of population and the true population has remained the same for the small area since the base year or the latest estimated year. For example, the population age 0 to 4 in a data zone may be estimated as the population aged 0 to 4 in the base year, times the Child Benefit count of 0 to 4 year olds for the current year divided by the Child Benefit count of 0 to 4 year olds in the base year. Special population groups (armed forces and prisoners) are removed from the population at the start of the process and updated counts of these groups are added towards the end of the process. This method is used by the Office of National Statistics (ONS) for the small area population estimates for England and Wales.

- 4.3 **Combination of both**, i.e. doing both and taking average – the combination method would involve creating a set of population estimates using the ratio change method and a set using the cohort component method. Once we have both sets of estimates the average of the two would be used to produce the final set of mid-year small area population estimates. This approach is currently used by the Northern Ireland Statistics and Research Agency (NISRA) for their small area population estimates.
- 4.4 **Combination of one of above plus administrative data sources** – Work began by colleagues in the Census Methodology team of the National Records of Scotland into using administrative data sources to produce population estimates at small area level. So a combination of either the cohort component or ratio change method could be used with the administrative data sources to produce a set of mid-year small area population estimates.

5. Reasoning

5.1 The following table highlights some of the advantages and disadvantages of each option.

	Advantages	Disadvantages
Cohort component method	Each individual part component is taken account of in the calculation. Most of the components we can be confident in our estimates, as provided from reliable sources.	The smaller the geographic level you are looking at the less reliable the migration estimates become, especially around the areas of high migration flows. (such as student areas)
Ratio Change method	For the ratio change there is no need to calculate migration. This takes away the element of the most uncertainty using the cohort component method.	The use of administrative data sources such as child benefit data or school census which are not currently available by single year of age at data zone level.
Combination Cohort Ratio	Following previous work carried out when reviewing the SAPE methodology it was discovered that the Cohort Component and Ratio Change methods are in general agreement for the majority of data zones. However on certain data zones the estimates appeared to diverge, the combination of both approaches may cancel out this divergence.	The combination approach has the same disadvantages as both the Cohort Component and Ratio change methods. Using migration which is most uncertain element of the cohort component method and using administrative data sources what are not currently available at single year of age or data zone level.
Combination with administrative data	Combining the administrative data sources to the SAPE methodology would be a good way to quality assure our estimates as we proceed.	At this current time, there are not enough administrative data sources available by single year of age or for the data zone boundaries to be able to produce a set of population estimates.

- 5.2 Due to the current lack of alternative administrative data sources available to us by single year of age and at the data zone geographic level the cohort component method continues to be the most viable method to use to produce the annual mid-year small area population estimates. We plan to investigate if we can make small changes to the council area methodology to improve the smaller area estimates going forward.
- 5.3 Following the exercise to redraw the data zone boundaries work has been carried out to see how many of the 2011 data zones are an exact match to the 2001 data zones. Looking at the physical boundaries it appears that only around 700 data zones are an exact one-to-one match. This number greatly increases if you look you do a one-to-one match based on population, to around 50 per cent of all data zones. However, due to the numbers of exact one-to-one matches by single year of age and sex and the additional time it would take to produce estimates matching those already produced it is recommended that this work is not done. Therefore there may be small differences between population estimates for the same year for those data zones where the 2001 and 2011 boundaries match.

6. Population estimates for geographies calculated as best fit from 2011 data zones

- 6.1 With the revision of the data zone boundaries, this gave the opportunity to evaluate the various geographies at which the population estimates are produced using a best fit approach from data zones. A paper has been produced comparing the fit of the 2011 data zone boundaries to the higher level geographies and will be published shortly on the NRS website.
- 6.2 Population Estimates best fit from 2011 data zones by Scottish Index of Multiple Deprivation (SIMD) will not be available until 2016 because a look up between 2011 data zones and the SIMD classification is not available.

Comments from PAMS on this approach are welcome.