Office for National Statistics

Moving to a new method for allocating postcodes to higher geographies: How does this affect births and deaths data?

December 2012

Editor: **N Richardson**

Office for National Statistics

2012

Office for National Statistics

1. Introduction

The National Statistics Postcode Lookup (NSPL) is a reference dataset published by ONS Geography that allows users to allocate data collected at postcode level to a range of higher geographies for the output of official and national statistics.

The methodology used by the NSPL to assign postcodes to higher geographic areas for statistical purposes has changed with the introduction of the 2011 output areas (2011 OA). This change means if the old and new methods were used to reference the same dataset (e.g. 2011 births and deaths) to the same geography (e.g. ward) the end results will not be identical.

In a time series this difference may be perceived as a change in the statistical data. However, the change could be as a result of the change in the method used to assign postcode data to higher geographies.

The first published edition of the new NSPL (2011 method) is the November 2012 edition, the old NSPL (2001 method) will still be released in parallel for a limited time.

This paper briefly explains the methods used and investigates the level of difference for 2011 births and deaths data for England and Wales referenced to 2011 Wards using each of the two NSPL assignment methods.

2. Changes to the NSPL

How has the method used to assign NSPL postcodes to higher geographies changed?

The structure of the NSPL and the way it is used by the end user has not changed. The unit postcode (e.g. PO15 5RR) is still used to reference user's data to higher geographic areas. This paper will use 2011 Wards as an example of a higher geographic area.

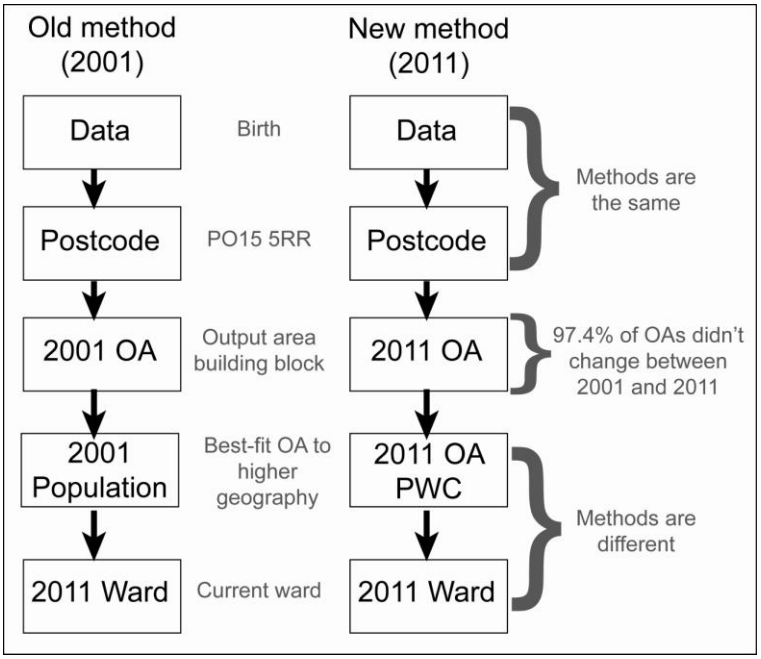
The method used by ONS to allocate each postcode on the NSPL to a higher geographic area has changed. Both involve assigning each postcode to an OA and then best fitting that OA to a higher geographic area (e.g. ward). However the set of OA building blocks has changed from the 2001 set to the 2011 OA set, where some 2.6 per cent of the 2001 OAs have changed as a result of the 2011 populations and so has the best-fit method. The two methods are explained below and figure 1 compares the methods.

Old method (2001): The postcode is assigned to the 2001 OA in which the mean address of the postcode (also known as its geometric centroid) falls (see NSPL User Guide for full details). ONS then assigns each OA to the ward in which the greater proportion of its 2001

Census population falls. Only ONS has access to the 2001 Census data, so only ONS can provide the OA to higher geography lookups.

New method (2011): The postcode is assigned to the 2011 OA in which its geometric centroid falls (see NSPL User Guide for details). Each OA is then assigned to the ward in which the OAs population weighted centroid (PWC) falls. The 2011 OA PWC is a single summary point that reflects the spatial distribution of the 2011 Census population in each instance of the 2011 OA. The PWCs for 2011 OAs have been published so are available for anyone to use.

Figure 1: The two methods for allocating postcode data to higher geographies.



3. Why might the two methods result in different allocations even if the data has not changed?

There are three reasons why a postcode may allocate to a different ward via the two methods:

A. **Output area change**. The 2011 OAs are the result of applying 2011 Census populations to the 2001 OAs. If the 2001 OAs were now too big in population size, they were split into two or more smaller 2011 OAs. If their population size had become too small, they were merged with one or more neighbouring OAs. Some OAs also changed as a result of re-alignment to changed LA boundaries. A fuller explanation of how OAs changed from 2001 to 2011 is here.

Where the OAs changed they will differ in size, shape and population altering the relationship between OAs and wards. As the NSPL has changed from using 2001 OAs to

2011 OAs as building blocks this could result in a postcode being assigned to a different ward in areas where OAs have been maintained. There are six whole local authorities where this factor does not apply because none or their OAs changed.

B. **Population change.** The population of an OA or the distribution of that population within the OA may have changed between 2001 and 2011. This change could result in an OA being allocated differently using 2001 or 2011 census data even if the same method and the same OA set were to be used.
Population change has the potential to affect the results of either method but the exact effect cannot easily be replicated. In practice, population change is a major reason for OA change (reason A above), however just because an OA has not changed does not mean its census population hasn't changed. The level of population change varies across England and Wales and is often (not always) reflected in maintenance of the OAs.
Only the new method uses the 2011 data, so any resulting difference in allocation should be for the better due to the more recent population data being used. It may be difficult to identify that a changed OA allocation is solely because of the effects of population change.

C. **Methodological Change**. In 2001 the OA was assigned to the ward containing the greater proportion of that OAs 2001 Census population. In 2011 a PWC was created for each OA from 2011 Census population data and then assigned to the ward within which it falls.
The different methods of assigning an OA to a ward will therefore sometimes provide different results.

4. Investigating the effect of the methodology change

Two versions of the NSPL were created using all postcodes available at October 2012:

- Postcodes to 2001 output areas (2001 OAs) using the old 2001 method to allocate a range of higher geographies (including 2011 Wards).
- Postcodes to 2011 output areas (2011 OAs) using the new 2011 best-fit method to allocate a range of higher geographies (including 2011 Wards).

Two datasets were available to be postcode matched:

- Births for 2011 by usual postcode of the mother for England and Wales residents only. This dataset included both live and still births.
- Deaths for 2011 by postcode of the deceased for England and Wales residents only.

The two datasets were referenced to each NSPL file in turn to add OA and ward codes to each birth or death. The results were compared to examine the effects of the change on the referenced data.

5. Results for births and deaths data

Both the births data and the deaths data matched successfully to the two NSPL files. All records received OA and ward allocations for both NSPL methods. Results were examined for differences in OA and ward allocations at an individual data level and then the data were aggregated to ward level and the counts examined for each ward to look for differences. The following sections explain the results.

a. All local authorities

Individual records

The allocations for each birth and death were compared and labelled according to whether the OA allocated was the same (i.e. unchanged between 2001 and 2011) or different (i.e. was changed between 2001 and 2011) and also whether the ward allocated was the same or different. As 97.4 per cent of the 2001 OAs still exist in 2011, these all remain comparable, so in these cases OA maintenance cannot explain a change in higher level allocation.

The majority of births (99.68%) and deaths (99.74%) have been allocated to the same ward by both methods. In table 1, the top two rows show records where the allocated ward was the same by both methods. Some of the remainder (i.e. Ward different but OA same) could be explained by either population change and/or methodological change. The remaining category is likely to be explained by the differing physical relationship between changed OAs and wards.

Table 1: Summary of allocations to individual records, count (per cent).

	Births		Deaths	
	count	**per cent**	**count**	**per cent**
Ward same & OA same	674,070	92.65	467,441	96.72
Ward same but OA different*	51,108	7.03	14,601	3.02
Ward different but OA same	1,177	0.16	917	0.19
Ward different & OA different	1,149	0.16	329	0.07
TOTAL	**727,504**	**100.00**	**483,288**	**100.00**

*If the OA is different, this means the OA was changed (e.g. split, merged or adjusted) between 2001 and 2011.

Office for National Statistics

Aggregate records (all local authorities)

Once wards have been allocated to the births and deaths they can be aggregated to produce a count for each ward. Ward counts for each method were compared and the results recorded in table 2.

The majority of wards have the same number of births (92.77%) and deaths (93.63%) whichever method was used to reference the data. Also of those that changed ward allocations most changed by 10 per cent or less.

Only a few wards (1.19% for deaths and 1.54% for births) changed by more than 10 per cent. The few wards not assigned, or assigned in only one of the two methods are, in the Isles of Scilly and City of London local authorities. Both local authorities are known for containing low population wards where some OAs are larger than wards.

Table 2: Comparison of birth/death counts by ward.

	Births: by ward		Deaths: by ward	
	ward count	**per cent**	**ward count**	**per cent**
Both methods produce the same birth/death count	7,967	92.77	8,041	93.63
Both methods provide zero birth/death count	15	0.17	20	0.23
Methods produce a different count (10% or less)	468	5.45	420	4.89
Methods produce a different count (more than 10%)	132	1.54	102	1.19
Data assigned in 2011 only (% change N/A)	3	0.03	2	0.02
Data assigned in 2001 only (Reported as -100% loss)	3	0.03	3	0.03
TOTAL	**8,588**	**100.00**	**8,588**	**100.00**

b. **Stable local authorities only (**Isles of Scilly, Allerdale, Eden, Christchurch, Craven and Worthing)

This subset of local authorities is unusual in that none of their OAs changed between 2001 and 2011. Looking at these areas in isolation means OA maintenance can't be a reason for any differing ward allocations from the NSPL. Only methodology and underlying population change are potential sources of ward allocation change in this subset.

Individual records (stable local authorities)

Almost all of births (99.92%) and deaths (99.91%) records have allocated to the same wards by the two methods. Table 3 shows that only 3 births and 1 death changed allocations when the new method was applied.

Table 3: Summary of allocations to individual records count (per cent) for the six stable LAs.

	Births		Deaths	
	count	**per cent**	**count**	**per cent**
Ward same & OA same	3,567	99.92	1,140	99.91
Ward same but OA different	0	0.00	0	0.00
Ward different but OA same	3	0.08	1	0.09
Ward different & OA different	0	0.00	0	0.00
TOTAL	**3,571**	**100.00**	**1,141**	**100.00**

Aggregate records (stable local authorities)

When aggregated by ward the births and deaths counts are the same in most of the wards (see table 4). The only real concerns are the two wards which only received data via a single method. Investigation has shown that this has occurred in the Isles of Scilly where two islands (Bryher and St Agnes) each have their own 2011 ward, however the two islands (and wards) are contained within a single OA. Only a single ward code must be assigned by the NSPL, in this case the 2001 methodology chose Bryher and the 2011 methodology chose St Agnes. As there are two postcodes on each island, whichever best-fit method is used, 2 postcodes will always be assigned the wrong way. The diagram in figure 2 illustrates this. In the births and deaths data two of the four records were accurately assigned in 2011.

Figure 2: Diagram showing the allocation swapping between islands in the Isles of Scilly.

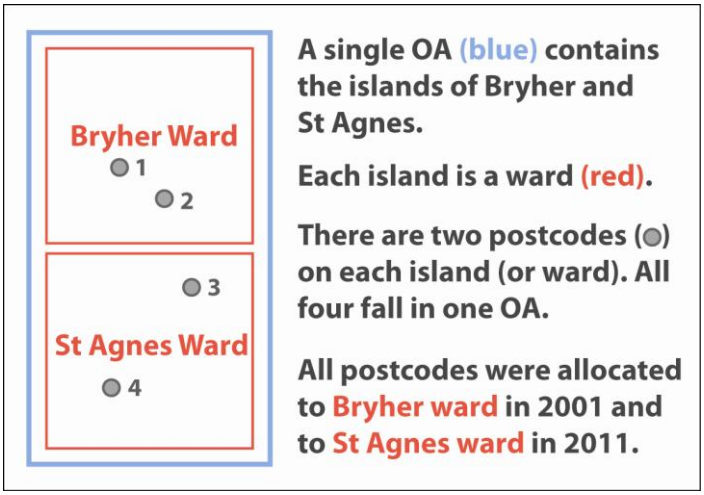


Table 4: Comparison of birth/death counts by ward (for stable LAs only).

	Births: by ward		Deaths: by ward	
	ward count	per cent	ward count	per cent
Both methods produce the same birth/death count	107	98.17	106	97.25
Both methods provide zero birth/death count	0	0.00	1	0.92
Methods produce a different count (10% or less)	0	0.00	0	0.00
Methods produce a different count (more than 10%)	0	0.00	0	0.00
Data assigned in 2011 only (% change N/A)	1	0.92	1	0.92
Data assigned in 2001 only (Reported as -100% loss)	1	0.92	1	0.92
TOTAL	109	**100.00**	109	**100.00**

6. Summary

On the evidence of using the new method NSPL for referencing the 2011 births and deaths, the effects of the change in methodology for assigning higher geographies to postcodes in the NSPL are small.

At individual record level, over 99 per cent of the data were allocated to the same ward when the methods were compared.

At the ward level over 92 per cent of the wards had the same counts between methods. Where the ward level data changed this was below 10 per cent in most cases. This figure accounts for all 3 reasons for change, not just methodological.

When only stable local authorities are taken into account (those not affected by OA change) almost all records have the same allocation. When aggregated, over 97 per cent of the ward counts are the same. The only changes observed were as a result of an unusual situation in the Isles of Scilly.