

Population And Migration Statistics (PAMS) Committee (Scotland)

Issues affecting the age distribution of the mid-year population estimates, 2002-2010 and 2012-2014

Introduction

1. National Records of Scotland (NRS) Population and Migration Statistics have recently become aware of errors in the mid-year population estimates for 2002 to 2014 because of an issue with an input data set used in the calculation of the mid-year estimates and unrelated processing errors.

September PNPM files

2. NRS Population and Migration Statistics receive monthly files, known as PNPM files, from NHS Central Register (NHSCR) containing estimates of moves to and from Scotland and between Scottish pre-April 2006 NHS Board areas by age (at mid-year) and sex derived from changes to General Practitioner (GP) registrations. These files are used to produce internal and cross-border migration estimates¹ for the year to mid-year. Furthermore, they are used to derive age-sex and geographical distributions for Office for National Statistics (ONS) Long-Term International Migration estimates of overseas migration.
3. For the revised mid-2002 population estimates (based on the 2011 Census) onwards NRS Population and Migration Statistics have assumed that there is a two month lag between a person moving and registering with a GP, meaning that registrations entered on to the NHSCR between 1 September and 31 August are assumed to correspond to moves that occurred within the mid-year period. Prior to this, for the unrevised estimates up to and including mid-2008 it was assumed that there was a three month lag between a person moving and registering with a GP, meaning that registrations entered on to the NHSCR between 1 October – 30 September were assumed to correspond to moves that occurred within the mid-year period. For the unrevised mid-2009 population estimates onwards the lag was changed to two months.

Figure 1: Calculation of age (at mid-year) from the PNPM specification

Age = Current Year – Year of birth
 If current month that job is run is > 9, 1 is added to age.
 If age is now not = zero:
 If month of birth > 6, 1 is deducted from age.
 If age > 99, age is set to 99.

Footnote

1) For the mid-2015 population estimates onwards, NRS Population and Migration Statistics will be using the NHSCR extract to produce internal and cross-border migration estimates. For more details refer to [PAMS \(15\) 10](#) on the NRS website.

4. The age (at mid-year) calculation in the PNPM specification uses date of birth (of the migrant) and processing date (that is the date that the job is run), as shown in Figure 1. The age calculation in the PNPM specification was not updated to reflect this change, when changes were made for the unrevised mid-2009 population estimates onwards to change the lag to two months instead of three months. The '9' in the second line of the code should have been replaced with an '8'. This change was not implemented and, therefore, all migrants in the September PNPM files for 2001 to 2013 (used to produce mid-year population estimates for 2002 to 2010 and 2012 to 2014) are one year younger than they were at the relevant mid-year as their age was calculated at the previous mid-year.
5. This issue primarily affects the age distribution of migrants. However, since migration estimates from the Community Health Index (CHI), derived by comparing annual CHI extracts one year apart, are scaled up (or down) to those derived from the PNPM files to produce migration flows for lower geographies (e.g. council areas and data zones), net migration and, consequently, population estimates for council areas and data zones may also be affected.
6. The mid-2011 population estimates rolled forward from the 2011 Census are not affected by this issue as the PNPM files for June to August 2011 were used to estimate migration in the period between Census Day (27/03/2011) and mid-2011 (30/06/2011).
7. For the mid-2015 population estimates onwards, this issue will be resolved since NRS Population and Migration Statistics will be using the NHSCR extract to produce internal and cross-border migration estimates (refer to [PAMS \(15\) 10](#)). Age at mid-year will be calculated from date of birth.
8. The impact of this issue on the mid-year population estimates for 2002 to 2010 will be partly resolved by the unattributable component of population change. The unattributable component of population change adjusted the rebased mid-year population estimates for 2002 to 2010 to ensure a smooth time series from 2001 to 2011. For more details refer to the methodology paper for the [Mid-2002 to Mid-2010 Revision](#) on the NRS website.

Impact

9. We have re-run the migration processing for 2012 to 2014 correcting for the issue with the September PNPM files highlighted in this paper. We have also taken the opportunity to correct for some other known minor issues with the mid-year population estimates. These include a correction to the age distribution of Scotland to England & Wales migrants, a correction to the lookup used to map CHI postcodes to higher areas in 2013 and 2014 and a correction to the sex-ratio of overseas out-migrants in 2012. The cumulative impact of these issues on the mid-2014 population estimates is discussed in the rest of this paper.
10. Table 1 shows the cumulative error in the mid-2014 population estimates by council area as a consequence of the issues outlined in this paper. Table 1 shows that the errors in the (total) council area populations are less than 0.10 per cent in all council areas. In percentage terms, the largest underestimate (positive error) is 0.06 per cent in Angus and the largest overestimate (negative error) is -0.10 per cent in Dundee City. In absolute terms, all errors are generally very small; the largest is an overestimate of 143 for Dundee City.

Table 1: Net cumulative error in the mid-2014 population estimates by council area ranked in descending order by the error as a percentage of the mid-2014 population

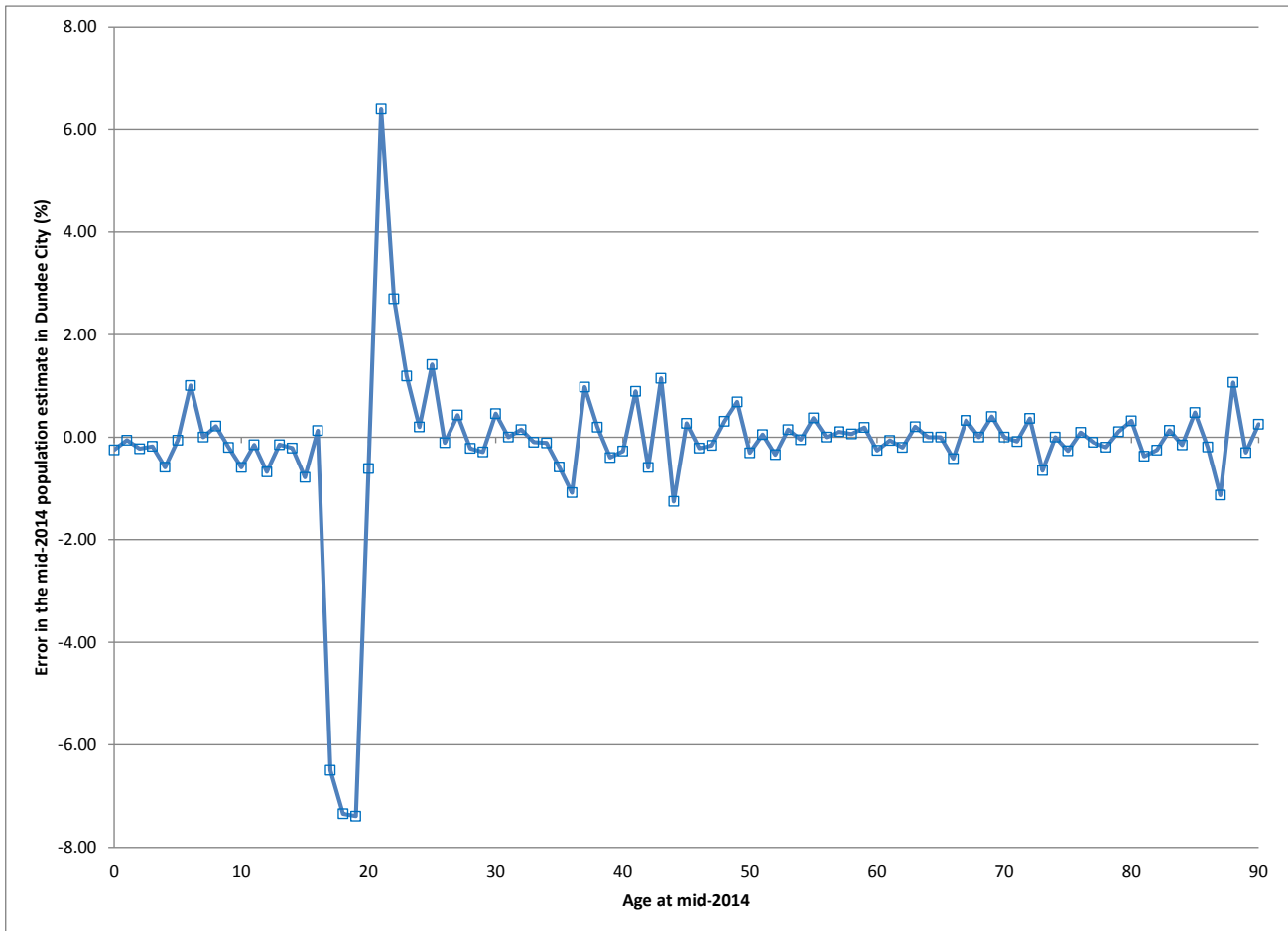
Council area	Mid-2014 population	Error ¹	Error ¹ as a percentage (%) of the mid-2014 population
Angus	116,660	74	0.06
Inverclyde	79,860	37	0.05
East Lothian	102,050	37	0.04
Falkirk	157,640	51	0.03
Perth & Kinross	148,880	43	0.03
West Lothian	177,150	41	0.02
South Ayrshire	112,510	25	0.02
Moray	94,750	19	0.02
North Ayrshire	136,450	26	0.02
North Lanarkshire	337,950	57	0.02
Eilean Siar	27,250	4	0.01
Aberdeenshire	260,500	33	0.01
Scottish Borders	114,030	13	0.01
Dumfries & Galloway	149,940	14	0.01
Midlothian	86,210	7	0.01
East Renfrewshire	92,380	6	0.01
Clackmannanshire	51,190	2	0.00
Renfrewshire	174,230	-1	0.00
Glasgow City	599,650	-4	0.00
Fife	367,260	-4	0.00
Orkney	21,590	-1	0.00
Highland	233,100	-16	-0.01
Shetland	23,230	-2	-0.01
East Ayrshire	122,150	-12	-0.01
Argyll & Bute	87,660	-10	-0.01
Edinburgh, city of	492,680	-76	-0.02
East Dunbartonshire	106,730	-18	-0.02
South Lanarkshire	315,360	-59	-0.02
West Dunbartonshire	89,730	-18	-0.02
Aberdeen City	228,990	-67	-0.03
Stirling	91,580	-58	-0.06
Dundee City	148,260	-143	-0.10

Notes:

1) A positive error means that the population has been underestimated and a negative error means that the population has been overestimated.

11. Although the errors in the total council area populations are small, the errors in the age distributions within council areas may be larger.
12. Figure 2 shows the error in the mid-2014 population estimates as percentage of the population estimate by single year of age in Dundee City.

Figure 2: Net cumulative error in the mid-2014 population estimates by single year of age in Dundee City



13. The largest overestimate in Dundee City is at age 19 (-7.4 per cent) and the largest underestimate is at age 21 (6.4 per cent). This is likely to be because of students moving to/from Dundee City to commence a course of Higher Education and then moving from/to Dundee City when they have completed their course. In general, students commence undergraduate Higher Education courses at ages 18 and 19 and Dundee City experiences net in-migration at these ages.

14. Table 2 shows the absolute error in the mid-2014 population by age and council area.

Table 2: Net cumulative error in the mid-2014 population estimates by age and council area

Council area	Age at mid-2014											
	All	< 17	17	18	19	20	21	22	23	24	25	> 25
Aberdeen City	-67	-45	-24	-166	-203	140	112	45	-23	-17	29	85
Aberdeenshire	33	-17	73	69	-19	-48	-66	5	14	-21	-6	49
Angus	74	-6	18	10	-18	-36	27	19	1	-1	50	10
Argyll & Bute	-10	-28	41	19	28	-41	-14	-18	-14	-9	6	20
Clackmannanshire	2	4	-1	4	-4	-2	13	13	-11	-6	7	-15
Dumfries & Galloway	14	-2	66	63	30	-99	-44	-16	-11	-9	18	18
Dundee City	-143	-40	-117	-160	-180	-17	191	93	40	6	39	2
East Ayrshire	-12	6	13	17	29	-31	-22	-8	-8	-6	13	-15
East Dunbartonshire	-18	5	14	31	8	-41	1	8	-1	-25	13	-31
East Lothian	37	-9	1	-44	28	-10	37	5	2	-23	0	50
East Renfrewshire	6	-3	28	32	5	-23	-31	-2	-10	-6	15	1
Edinburgh, City of	-76	-35	-328	-822	-436	361	555	-70	-91	46	346	398
Eilean Siar	4	-8	8	16	-4	-8	-14	2	6	-11	3	14
Falkirk	51	1	-14	-19	23	9	37	-11	21	-8	24	-12
Fife	-4	-11	-147	-422	-67	401	206	-84	-21	-37	53	125
Glasgow City	-4	-21	-159	-173	-229	-4	224	0	-30	65	79	244
Highland	-16	-4	123	87	59	-131	-119	-31	-25	38	-12	-1
Inverclyde	37	-5	16	6	13	-18	-10	-14	12	14	-22	45
Midlothian	7	-14	-3	-9	11	10	-6	-2	1	-3	9	13
Moray	19	1	32	20	1	-7	-36	7	-19	5	-5	20
North Ayrshire	26	-1	31	18	10	-59	-3	8	10	-3	-30	45
North Lanarkshire	57	15	29	37	18	-35	-31	-4	-3	-12	-22	65
Orkney Islands	-1	3	5	14	-7	5	-8	-11	0	-13	14	-3
Perth & Kinross	43	-3	30	-11	17	-48	-6	49	-23	37	-15	16
Renfrewshire	-1	5	51	3	12	-29	-11	-5	-32	29	8	-32
Scottish Borders	13	-5	22	40	-13	-9	-55	4	-9	4	-3	37
Shetland Islands	-2	-3	11	11	8	-4	-15	-14	0	-2	-4	10
South Ayrshire	25	4	32	31	10	-21	-34	-28	10	15	5	1
South Lanarkshire	-59	-15	14	44	30	-39	-42	-1	-14	-16	20	-40
Stirling	-58	1	-120	-172	-21	22	113	30	15	-7	20	61
West Dunbartonshire	-18	-1	1	-7	-3	-13	-5	-4	12	5	9	-12
West Lothian	41	3	-2	14	-21	21	-24	3	0	9	19	19
All	0	-228	-256	-1,419	-885	196	920	-32	-201	38	680	1,187

Notes:

A positive error means that the population has been underestimated and a negative error means that the population has been overestimated.

Proposal

15. The issues highlighted in this paper:

- do not affect the estimate of Scotland's total population,
- do not affect the mid-2011 population estimates,
- are partly resolved by the unattributable component of population change in the rebased mid-year population estimates for 2002 to 2010,
- have only a small effect on total population estimates for council/NHS Board areas, and
- affect primarily the age distribution of the population in the age range 17 to 25.

16. Given that most of the branch's publications use the mid-year population estimates, e.g. the national and subnational population projections, small area population estimates and life expectancy estimates, if we were to revise the mid-year population estimates we would also need to revise all of the publications which use them. This would require a lot of additional resource and would impact on ongoing development work including the review of the subnational population projections and the project to include Higher Education Statistics Agency (HESA) data in the migration estimates. This seems excessive given that the issues mainly affect the age distribution of the population at student ages.
17. Given the uncertainty that is attached to population estimates the magnitude of the cumulative error described here is probably within these bounds.
18. NRS Population and Migration Statistics, therefore, propose to resolve these issues going forward but make
 - no correction to the mid-year population estimates for 2002-2010 and
 - no correction to the mid-year population estimates for 2012-2014 until these estimates are rebased to better reflect the results of the 2021 Census.
19. If this approach is adopted, it should be emphasised that we will be carrying forward errors in the population and migration estimates for earlier years which will affect the population estimates for mid-2015/2014-based population projections onwards. The errors in the mid-year population estimates for 2012-2021 will be fully resolved only when these estimates are rebased to better reflect the results of the 2021 Census.
20. We will document the impact of the error and publish it on the NRS website.
21. Back in 2009 PAMS discussed the criteria for revising population estimates and [the minutes from the meeting and the paper are on the NRS website](#). A view shared by most at the meeting was that revisions create a big workload for Local Authorities and that they should be kept to a minimum.
22. The views of PAMS members on this proposal are welcome.

NRS: Population and Migration Statistics
18 September 2015