

Tending Unattributable Population Change

Further information at the request of the Inspector

As discussed at the EiP matter 3 session, Neil McDonald has been working with the help of ONS statisticians¹ and John Hollis, to better understand the 'Unattributable Population Change' error in Tending.

The Inspector requested further clarification regarding correspondence between the ONS and Neil McDonald (on behalf of Tending). This confirms there is a significant UPC error here. The Inspector also asked for conformation of the scale of UPC in the wider Housing Market Area.

In **Appendix 1** ("Thoughts on Issues") the ONS say that 'at most 47%-57%' of the error relates to misreported migration.

PBA understands that this was provided as an email attachment between Mr McDonald and Neil Park, Senior Statistician in the ONS Population Estimates Unit in summer 2017. This is also referenced in several participants hearing statements.

In the second email [**Appendix 2**] (November 2017), the ONS agree that the migration error is likely to be in the range 5-6,000 which equates to 47-57% of UPC i.e. that this is not a maximum. They then go on to say that it is difficult to conceive how there could be any more than 4,500-5,000 error due to the censuses.

Using this second email Mr McDonald has concluded that if the maximum census error is 5,000 (equating to 47% of UPC) then it follows that the minimum migration error is 53%. This suggests that the ONS believe that the error due to migration is more likely to lie in the top half of their range of 47-57% i.e. in the narrower range 53%-57%. Mr McDonald uses this evidence to conclude that the ONS have, in effect, confirmed his conclusion that the UPC error is most likely to be 55%, the middle of a very narrow range (53%-57%)

[Note: part of this email is redacted – this is because it relates to an informal comment not intended for public consumption on about the progress of a separate piece of research.]

¹ Neil Park; Senior Statistician in the ONS Population Estimates Unit.



Tending Unattributable Population Change

At **Appendix 3**, in response to the Inspectors query regarding UPC in the wider HMA, we provide a brief analysis which demonstrates the scale of UPC in the HMA. This also shows that the adjustment in Tending does not require other adjustments elsewhere. This is an extract of various reports prepared by Neil McDonald as part of ongoing appeals in the district.

Tendring Unattributable Population Change

APPENDIX 1

NOTES BY THE ONS ON TENDRING'S POPULATION ESTIMATES

Thoughts on issues with the population estimates for Tendring between 2001 and 2011

1. Population estimate rolled forward from 2001 was 10,533 higher than the 2011 Census based population estimate; the rolled forward estimates 7.6% higher than the Census based estimate.
2. Assume Births and deaths are fine
3. Assume any negative impact of international migration is low due to there being few international migration moves (5,197 in, 3732 out over the decade, net=1465).
4. Internal migration has a potentially larger impact than the other components due to the volume of moves (65,284 in, 48,404 out, net=16,880). Our traditional view of this has been that areas that gain population over a period are more prone to underestimation than overestimation on the basis that.
 - a. Assuming human behaviour is relatively constant we miss moves at a constant rate both into and out of any location. Rates of missingness will vary by age and sex but should vary little for inflows and outflows.
 - b. On a net inflow we will miss more moves, in absolute terms, on the inflow than the outflow.

For an area like Tendring, with net inward internal migration, we'd probably end up underestimating as a consequence. But, Tendring tends to have net internal migration of relatively well behaved people in their middle/old age, we tend not to miss moves for these age groups as they interact well with GPs.

Your conclusion, 5.6, is difficult to reconcile with some of this logic. If we are missing flows out of Tendring, measured primarily via GP registrations it would follow that we must also be missing some flows into Tendring as these are measured via the same method. Given that flows in are in excess of flows out we would expect to miss more flows into Tendring than flows out of Tendring.

5. The relationship between the PR and mid-year estimate/Census in 2001 provides us with two possible avenues for believing that the 2001 Census may have overstated the population – or more accurately - how the 2001 Census and 2001 PR may have lead to overestimation of the 2011 MYE.

Tendring Unattributable Population Change

The first is relatively obvious; the 2001 Census sits above all of our comparator admin data (mainly the PR and state pension's recipients) for a large number of age groups. The 2001 One Number Census QA pack for Tendring shows this <http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/quality-of-the-census-data/local-authority-data-quality/england/s-t/tendring-onc-qa-pack.pdf>. The charts at the end of the document show Tendring had an unusual relationship between the PR and the Census in 2001; generally the PR sits above the Census estimates, in Tendring the PR sits below the Census for the majority of age/sex groups above 45. These are the charts I supplied via email previously.

The second is a bit more complicated and is not actually about the 2001 estimate being overestimated. If we imagine that the 2001 census for Tendring was perfect, this means that the patient register is missing large numbers of people. We drive our internal migration estimates using data from the patient register, any moves involving people resident in Tendring in 2001 (and captured/estimated by the Census) but absent from the 2001 patient register will have been missed. Therefore we will miss outflows from Tendring, and the most likely consequence of this is that we will overestimate the population. I suspect the first case (the Census being overestimated) to be more likely than the second case (the PR understating the population) for Tendring. The second scenario tends to occur in inner London local authorities with very high levels of internal migration for 20-40 year olds who are less likely to interact regularly with GPs.

Further, the 95% confidence intervals give us a range within which we would expect the population estimate to fall 95 times out of 100. 5 times out a 100 the estimates could fall outside of this range. The confidence interval around the Census estimates should not be taken as a guarantee that an estimates is within a particular range. Following the 2001 Census a number of adjustments were made to the mid-year population estimates to account for inaccuracies in the 2001 Census. Some of the issues with the 2001 Census are discussed in this paper <https://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/local-authority-population-studies/2001-census---local-authority->

Tending Unattributable Population Change

[population-studies--full-report.pdf](#). The 2011 Census learnt from a lot of the difficulties of the 2001 Census and produced more robust population estimates as a result.

6. The age/sex distribution of the discrepancy for Tending is quite informative. Generally speaking the discrepancies between Census based and rolled forward estimates are greatest for the young adult population. This reflects the high level of population churn for this group and the difficulty in measuring internal migration for this group given their generally good health and their poor levels of interaction with the health service. Tending's discrepancies are different, they tend to be spread relatively evenly a wide part of the age range with a particularly large amount of discrepancy at the end of the age distribution. Given the low volume of moves for older people (say 70+), due to both internal migration and international migration, for these individuals it is difficult to see how this error could have been caused by migration. Much more likely is that we started off with a base population that was slightly overestimated and this discrepancy was carried through the entire decade.

It is also interesting that the discrepancy is relatively symmetrical for males and females (5,682 males, 4,851 females). If the problem were overwhelmingly internal migration based we might also expect the discrepancy to be substantially larger for males than females.

7. As you may be aware we are in the process of changing some of methods ([see appendix 2 for details](#)). One of these changes involves removing part of internal migration process called "scaling factors", these were used to adjust the level of raw internal migration flows picked up using the patient register and HESA to account for moves by people who did not appear on the beginning and end patient register (those who were born, died, immigrated or emigrated during 12 months preceding the mid-year point) and those who moved more than once during the year. In effect this applied a multiplier to the levels of in inflows and outflows to each local authority. The removal of these scaling factors has, very recently, revealed some interesting side effects of the scaling process. For Tending the impact of scaling factors would have been to make internal migration flows increasingly positive and may have been a partial contributor to the unattributable difference found in 2011.
8. Unfortunately, I'm not an expert on housing statistics but I know that the analysis was subject to a high degree of quality assurance.

Tending Unattributable Population Change

9. Regarding your conclusions. I agree with 5.1, 5.2 and 5.5.

On 5.3 we think the evidence suggests that around 4,500 of the discrepancy is due to the 2001 Census base. Some of the remaining difference may be due to sampling error relating to the 2011 Census but this is still likely to leave 5-6,000 of the difference unexplained.

On 5.4 we would therefore think that the discrepancy due to migration is likely to at most 5-6,000. The impact of scaling that I mentioned in point 7 *may* account for 3-4,000 of the difference. Our traditional viewpoint on LAs such with net internal inflows has been that any internal migration discrepancy would have lead to an underestimate of the population rather than an overestimate.

On 5.6 I've mentioned, as part of 7, that because we don't simply use GP patient registrations, it is possible for us to both over-estimate the inflows and the outflows. The findings I mentioned in point 7 are only about 2 weeks old, assuming these don't change (a flaw in the analysis could yet be found) it suggests that overestimation of the inflow was a more significant driver of the discrepancy than underestimation of the outflow.

This would leave around 2-3,000 of the difference unexplained. The "[understanding discrepancies tool](#)" I previously linked you to suggests that international immigration may have been overestimated for young these are the charts I included (I've added the 2011 equivalents as well).

Tendring Unattributable Population Change

APPENDIX 2

From: Park, Neil [mailto:neil.park@ons.gov.uk]
Sent: 29 November 2017 13:43
To: Neil McDonald <neilmcdonald@googlemail.com>
Subject: RE: Tendring's UPC

Hi Neil,

Apologies for not replying sooner and thank you being part of the Household Projections Collaborative group meeting last week.

Attached is data for Tendring's GP list inflation in 2001 (this is from the [One number Census QA pack for Tendring](#)) and a link to the [2011 Census QA pack](#). Apologies for doing this in 'kit-form' but it makes the process of getting data to you a bit more simple for me.



In working out the potential error attributable to each component of the MYEs in the previous decade we have had to be quite speculative - we do not have any way of calculating the actual amounts attributable to each component, all we've got is an understanding of how the methods work and some logical deductive reasoning. I don't think you would be over-simplifying to say that the migration error is likely to be in the range 5-6,000. This is partly because it is difficult to conceive how there could be any more than 4,500-5,000 error due to the 2001 and 2011 censuses and practically no other areas of the process that could cause error.

regards,
Neil

[Note: part of this email is redacted – this is because it relates to an informal comment not intended for public consumption on about the progress of a separate piece of research.]

Tending Unattributable Population Change

APPENDIX 3

Implications for the rest of the Housing Market Area

- 1.1. Given that UPC across England as a whole sums to a negligible amount it is natural to ask whether errors in over-estimating net migration flows into Tending have resulted in the under-estimation of flows into other parts of the HMA and, most particularly, into Tending's neighbour, Colchester. However, an examination of the detailed data for UPC across the HMA shows that this is not the case.
- 1.2. The figure below compares the scale of UPC across the HMA both as a percentage of the census population in 2011 and as a proportion of the population change between 2001 and 2011. The enormous figure of 1423% as the proportion which Tending's UPC is of the population change between the 2001 and 2011 censuses is the result of UPC being 10,533 when the census-based estimate of population change was a fall of 740 people ($10,533 \div 740 = 1423\%$).

Figure 6.1: Comparison of UPC across the HMA

	Colchester	Chelmsford	Braintree	Tending
UPC as % of pop in 2011	-2.4%	0.6%	0.7%	-7.6%
UPC as % of pop change	-23%	9%	7%	1423%

- 1.3. As can be seen, UPC in Chelmsford and Braintree is small. Colchester has moderate UPC but it is very much smaller than in Tending. Moreover, it is negative – like Tending. Had migration flows been misallocated to Tending instead of Colchester you would expect to see UPC of the opposite sign in Colchester.
- 1.4. This lack of any offsetting UPC in Colchester is confirmed if the age profile of UPC in the two authorities is compared as shown in the figure below.

Tending Unattributable Population Change

