Office for National Statistics

Date: 30 April 2013
Coverage: England and Wales
Theme: Population

Methods used to revise the subnational population estimates for mid-2002 to mid-2010

1. Introduction

This paper forms part of the release of the revised mid-2002 to mid-2010 subnational population estimates for England and Wales. The revised estimates go down to local authority (LA) level and create a continuous series between the mid-2001 and the mid-2011 population estimates.

This paper summarises the methods used to create the revised estimates, and also includes case studies of three LAs – Newham, Luton and Manchester – to demonstrate how the methods have worked in practice. The case studies begin on page 14. All data and reports referred to are available via the 'References and related publications' section on page 23.

2. Background

The population estimates for mid-2002 to mid-2010 have been revised to bring them into line with the official mid-2011 estimates, which are based on the 2011 Census (27 March 2011) estimates of the usually resident population, plus the effect of births, deaths and migration up to 30 June 2011.

The former mid-2002 to mid-2010 estimates were based on change as a result of published data on births, deaths and migration since mid-2001. If the 2011 Census estimates had not been available the equivalent mid-2011 population estimates for each LA would have been different: in some cases lower, in other cases higher. More analysis is provided in the statistical bulletin accompanying this release (available via the link above).

This paper describes how the potential causes of these differences have been determined, and how they have been distributed across the decade to create the revised mid-2002 to mid-2010 series.

These revisions should be seen in the context of the revised national (England and Wales) mid-2002 to mid-2010 estimates, which were published on 13 December 2012.

3. Definitions and principles

The process of creating annual population estimates through the addition of the effects of births, deaths and migration is referred to as 'rolling forward'. In this paper the estimates rolled forward from mid-2001 are referred to as the 'rolled-forward mid-2011 estimates', and those created using the 2011 Census estimates as the starting point are referred to as the official (Census-based) mid-2011 estimates.

All annual population estimates are based on the usual resident population, which excludes the effects of short-term international migration. This means that immigrants who do not intend to stay in the UK for at least 12 months are excluded. Conversely, emigrants who intend to stay outside of the UK for less than 12 months are included.

The factors contributing to the difference between the rolled-forward and official mid-2011 estimates are referred to in this paper as 'components of difference'.

A key principle of the national mid-2002 to mid-2010 revisions was that components of difference would initially be identified at national level, as they could not reliably be derived directly at subnational level. However, it is also essential that the components of difference at LA level sum to the national totals. This can mean one of two things:

1. Components of difference which existed in the national revisions have been broken down to subnational level with their original national total being retained.
2. Components of difference which only exist at subnational level (and so did not feature in the national revisions) have been processed so that their national sum is zero.

For each component methods have been applied consistently across the country. However, exceptions to this principle have been applied to two LAs where specific issues were identified: Harrogate, and Oadby and Wigston. These are explained in more detail in due course.

In practice there is substantial estimation around most of the identified components of difference. Instead they should be considered as best approximations, based on the evidence available, of how the difference between the rolled-forward and official mid-2011 estimates arose.

As with the national revisions, once identifiable differences have been taken into account the remaining difference for each LA has been allocated to a general 'Other' component rather than being arbitrarily, and potentially incorrectly, assigned to specific causes. This approach is in line with international best practice.

4. Summary of components of difference

The official estimate of the mid-2011 population of England and Wales is 56,171,000, which is 464,000 higher than the rolled-forward mid-2011 estimate of 55,707,000.

The revised national population estimates for mid-2002 to mid-2010 attributed this 464,000 difference as shown in Table 1 (overleaf). A detailed explanation of what is meant by each of these components and how they were derived was presented in the report 'Methods used to revise the national population estimates for mid-2002 to mid-2010', which accompanied the release of the revised national estimates in December 2012.

Table 1: Components explaining the difference between the mid-2011 official and mid-2011 rolled-forward estimates for England and Wales: summary

Components	Impact on difference	Remainder
Initial difference	n/a	464,200
EU8 immigration adjustment	250,000	214,200
Republic of Ireland migration roll-back	65,800	148,500
Migrant switcher roll-back	37,000	111,500
Visitor switcher roll-back	-7,500	119,000
Armed forces adjustment	-7,100	126,100
Cross-border migration correction	2,400	123,700
Mid-2009 asylum seekers and visitor switchers correction	- 11,600	135,300
Removal of historic processing adjustments	800	134,500
Other	134,500	0

Note: totals may not sum because of rounding.

This report summarises these national components and explains how they have been broken down to LA level. It also explains a number of components that did not affect the national total but did have an impact at LA level:

1. Replacement of international immigration flows for the years ending mid-2006 to mid-2011 with flows based on a new method developed as part of ONS's Migration Statistics Improvement Programme. This change also impacted on the international emigration values.
2. Changes to the way in which specific adjustments relating to the Home Armed Forces (HAF) have been distributed.
3. Removal of the separate school boarders component from the mid-year estimates.
4. Adjustments to correct for specific issues identified in two LAs: Harrogate, and Oadby and Wigston.

This paper details the methods used for each component and then provides case studies of the effect of the various components of difference in selected LAs.

5. Note on the impact of rounding and constraining

When the revised national estimates for mid-2002 to mid-2010 were created, some components of difference were affected by rounding and were not constrained back to the totals shown in Table 1. This meant that the published estimates had components of difference which were very slightly different. For practical purposes these differences are inconsequential and this paper refers throughout to the Table 1 values. However, users who download the components of difference file will be aware of the small differences.

For the subnational processing, in most cases the methods used to derive the LA-level components of difference initially led to non-integer values (values that are not a whole number), both for LA totals and the individual age/sex groups within them. For each component these were then rounded and constrained back to the separate totals for England and for Wales from the revised national back series.

This rounding and constraining process will have led to small changes to individual age / sex / LA values. However, the effect is minor. Furthermore, given the uncertainty in the estimation process, the true 'real world' values in each of these cases cannot be known.

6. International migration

There are several aspects of the revisions to international migration values at LA level:

i) Switch to the new method immigration breakdown

The previous series of subnational estimates used a modelling approach to determine the number of long-term international immigrants. This approach took national and regional immigration totals from the International Passenger Survey (IPS) and used a range of demographic, social and economic measures to model how many were likely to have gone to each LA.

In November 2011, as part of its Migration Statistics Improvement Programme, ONS published immigration totals using an improved method for the years ending mid-2006 onwards. At the time the new method totals were described as 'indicative' and did not replace the official totals created using the old method. However, following analysis and user consultation, the new method has been adopted as the way forward for future estimates.

In addition the new method (indicative) estimates for the years ending mid-2006 to mid-2011 have been applied to the revised estimates, replacing the previous international immigration figures. However, for the years ending mid-2002 to mid-2005 the existing immigration flows have been retained as the administrative data required by the new method are not available.

The 'References' section of this report contains links to further information on both old and new methods.

ii) Impact of the new immigration breakdown on emigration figures

Estimates of long-term international emigration from each LA are also created using a modelling process. As the international immigration figures are used in the model, the model has been re-run with the new immigration data for the years ending mid-2006 to mid-2011.

The emigration values from the model are also used to inform calculation of the number of emigrant visitor switchers (people whom the IPS data count as originally intending to be outside the UK for less than 12 months, but who actually stay abroad longer and so become long-term emigrants). Therefore the LA-level emigrant visitor switcher figures for the years from mid-2006 onwards have also been revised.

iii) Additional EU8 immigration

Summary of national adjustment

The revised national back series included 250,000 additional long-term immigrants from the EU8, the eight countries of central and eastern Europe that joined the European Union in 2004: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia. These 250,000 were immigrants who were believed to have been missed from the original immigration estimates (see national back series methods paper for explanation). They were distributed over time as follows:

Table 2: Distribution over time of additional 250,000 EU8 immigrants to England and Wales

Year ending	Additional EU8 immigrants
Mid-2005	40,000
Mid-2006	40,000
Mid-2007	56,000
Mid-2008	56,000
Mid-2009	35,000
Mid-2010	13,000
Mid-2011	10,000
Total	250,000

Note: the addition of migrants to the year ending mid-2011 only covers the period up to Census Day.

LA-level breakdown for years ending mid-2005 to mid-2010

The additional flows for the years ending mid-2005 to mid-2010 have been broken down to LA level using the distribution of workers with a European Union (EU) nationality other than British.

The distribution has been derived from the new immigration breakdown methods described above. Part of the new methods involves constraining administrative sources data on the number of immigrant workers in each LA to 'subcontinental' worker totals from the IPS, with the EU (specifically all EU nationalities apart from British) forming one subcontinent grouping.

The reason why the additional EU8 migrants have been distributed using the overall EU grouping rather than an EU8-specific distribution is that all non-British EU workers – both EU8 and other EU – have been suppressed proportionately by the constraining process. Therefore if the missing workers are added using the geographic distribution of all EU workers (regardless of where in the EU they actually originate), the true numbers for both EU8 and other EU workers are restored.

This concept can be difficult to visualise, so here is a worked example based on two fictitious LAs:

i. Suppose that nationally in any one year there are 300,000 EU immigrant workers. 200,000 of these are included in the IPS-based totals, and 100,000 are EU8 workers who have been missed.

ii. Also suppose (for simplicity) that all of the EU8 workers who came into the country were missed.

iii. The new immigration method distribution suggests that 1% of EU workers go to Borsetshire, and another 1% to Sodor. This means that in the existing (IPS-based) totals each area is allocated 2,000 EU workers, as opposed to the 3,000 each area would have had if all EU workers had been included.

iv. Suppose too that all of Borsetshire's EU immigrant workers are from EU8 countries, but all of Sodor's come from the rest of the EU. This means that in the IPS-based totals (where all EU8 immigrants are missing), Borsetshire has wrongly been allocated 2,000 immigrants originating in the rest of the EU, whereas Sodor has been given 1,000 too few.

v. Applying an EU8-specific distribution (if known) to the missed 100,000 EU8 migrants would not be appropriate. It would add 3,000 to Borsetshire (giving a total of 5,000) and none to Sodor (leaving the total at 2,000).

vi. However, if the EU worker distribution were used, it would add 1,000 to both Borsetshire and Sodor, giving the correct total of 3,000 in each of them.

In the real immigration estimates only a proportion of EU8 workers have been missed each year, and it is unlikely that there are any LAs where all EU workers are of EU8 nationality. But the principle holds: using the EU worker distribution is the best option.

In practice not all missed EU8 immigrants are workers. However, most EU8 immigrants do work and, as developing specific methods for non-workers would be complex and subject to considerable uncertainty, the EU worker approach has been applied to all 240,000 additional EU8 immigrants in the years ending mid-2005 to mid-2010.

Because of the small sample size in the part of the administrative data which determines the split by subcontinent at LA level, using an EU worker distribution for each individual year would cause substantial year-on-year variation and is not considered robust. Therefore a single five-year average distribution has been applied to all six years, based on the combined EU worker distributions for the years ending mid-2006 to mid-2010 (no EU worker distribution is available for the year ending mid-2005).

LA-level breakdown for year ending mid-2011

The new method processing for the year ending mid-2011 did not constrain to subcontinent totals but instead constrained just to the full ‘all immigrant’ national IPS estimate. Therefore, using a similar reasoning to the Borsetshire / Sodor example above, the 10,000 EU8 workers added in that year have been distributed to LA level using the all immigrant distribution.

Transfer from England to Wales

The separate totals for England and for Wales in the revised national back series were derived from the new method all immigrant distribution as it had not been identified at that stage that using the EU worker distribution was a better means to determine the LA values for the years ending mid-2005 to mid-2010. The need to match those published national totals has had a small effect on LA totals, with a net transfer of around 1,000 from England

to Wales. However, this is a minor adjustment which is considered acceptable given the degree of uncertainty around the true values.

Age/sex breakdown

The revised national back series split the additional EU8 immigrants by age and sex using the national distributions from the new method all immigrant estimates (with the values for the year ending mid-2006 also being applied to the year ending mid-2005). All LAs have been assumed to share the national age/sex distribution for each year.

iv) Republic of Ireland migration roll-back

The revised national back series included an increased net flow between the Republic of Ireland and the UK of 65,800, comprising 23,450 more immigrants and 42,350 fewer emigrants. This increase was spread evenly across the seven years between the year ending mid-2002 and the year ending mid-2008. The justification for this was the roll-back of a new method which had originally only been applied from the year ending mid-2009 onwards.

In the former back series the flows between the Republic of Ireland and the UK were processed separately from other international migration flows. This means that they had their own specific LA-level distributions. In the revised back series these distributions have been applied to the emigration flows for all years and the immigration flows up to the year ending mid-2005.

As the Republic of Ireland is part of the EU then for the additional immigrants in the years ending mid-2006 to mid-2008 the new method EU worker distribution has been applied, on the same logic that it was applied to the additional EU8 workers. In this case a single three-year average distribution has been applied based on the combined EU worker distributions for the years ending mid-2006 to mid-2008.

The age/sex breakdown of the revised flows to and from the Republic of Ireland matches the distribution in the former flows for each year, with the same age/sex distribution in all LAs.

v) Migrant switcher roll-back

The International Passenger Survey (IPS) asks people how long they intend to stay in their country of destination (either in the UK for arrivals, or out of the UK for departures). If they are staying 12 months or more they are regarded as a migrant; if they are staying less than 12 months they are regarded as a visitor.

However, some people who originally intend to stay for 12 months or more (and so are initially considered to be migrants) actually stay for less than 12 months and so are re-classed as visitors. Such people are termed ‘migrant switchers’ and ONS migration estimates include adjustments for this.

In the former back series the methods for estimating the number of migrant switchers had only been applied to the estimates from the start of 2004 onwards. The revised national estimates for mid-2002 to mid-2004 extrapolated the methods to cover the full decade by assuming that the change caused by the new methods during the first six months of 2004 would be representative of the rate of change in the first part of the decade. This assumption led to the changes shown in Table 3.

The overall effect was that migrant switcher net flows were reduced by 37,000, with the result that net international migration was increased by 37,000.

Table 3: Changes to England and Wales migrant switcher estimates due to roll-back of new methods

Year ending	Inflows	Outflows	Netflows
Mid-2002	-6,800	8,000	-14,800
Mid-2003	-6,800	8,000	-14,800
Mid-2004	-3,400	4,000	-7,400
Total	-17,000	20,000	-37,000

The changes to the migrant switcher flows have been broken down by LA, age and sex using the same distributions as the former flows for each year.

vi) Visitor switcher roll-back

Visitor switchers are people who originally intend to stay for less than 12 months but then go on to stay for 12 months or more (either in the UK for arrivals, or out of the UK for departures). This means that they are re-classed from visitors to migrants. Such people are termed 'visitor switchers' and ONS migration estimates also include adjustments for this.

As with migrant switchers, in the former back series the methods for estimating the number of visitor switchers had only been applied to the estimates from the start of 2004 onwards. The revised national estimates for mid-2002 to mid-2004 extrapolated the methods to cover the full decade by assuming that the change caused by the new methods during the first six months of 2004 would be representative of the rate of change in the first part of the decade. This assumption led to the changes shown in Table 4.

The overall effect of this is that visitor switcher net flows were reduced by 7,500, with the result that net international migration was also reduced by 7,500.

Table 4: Changes to England and Wales visitor switcher estimates due to roll-back of new methods

Year ending	Inflows	Outflows	Netflows
Mid-2002	-10,600	-7,600	-3,000
Mid-2003	-10,600	-7,600	-3,000
Mid-2004	-5,300	-3,800	-1,500
Total	-26,500	-19,000	-7,500

The changes to the visitor switcher flows have been broken down by LA, age and sex using the same distributions as the former flows for each year.

vii) Mid-2009 asylum seekers and visitor switchers correction

The revised national estimates for mid-2009 included corrections to processing errors in the former values for both asylum seekers and visitor switchers. The result of this was:

- 300 additional asylum seeker immigrants,
- 600 additional asylum seeker emigrants, and
- 11,400 additional visitor switcher emigrants.

The net effect of the correction was a reduction of the mid-2009 population estimate by 11,600, which fed through to an identical reduction in the mid-2010 estimate.

The corrected asylum seeker totals were run through the standard mid-year estimates processing for distribution by LA, age and sex, so all existing values have been replaced by the new values.

The corrected visitor switcher totals were used in the re-processing of visitor switcher emigrants referred to in Section 6ii, so no separate action was required.

viii) Note on processing error

A processing error has been identified in the revised national population estimates for mid-2002 to mid-2010. This affected the age/sex distribution of visitor switcher immigrants to England in the year ending mid-2003, and therefore the population estimates for mid-2003 and subsequent years up to mid-2010.The maximum error for any age/sex group was 1,300 (0.4% of the total) for females aged 17 in the estimate for mid-2003, with the error reducing through the decade to zero in mid-2011.

As the error is in the sex/age distribution, the total population estimates for each year have not been affected. The impact of the error has been assessed as minor and, in accordance with ONS's Population Statistics Revisions Policy, the published estimates will not be revised to correct the error. The revised subnational estimates have been designed to be consistent with the national back series, and so also include the error. However, the effect on any one local authority is negligible.

7. Armed forces

i) Home Armed Forces (HAF)

National armed forces adjustment

Table 1 indicates an 'Armed forces adjustment' of -7,100 in the revised national back series. This was partly due to a correction of processing errors in the former estimates, which have also been corrected at LA level. However, the main driver of the adjustment was a change in the method for the annual counter-adjustment to the civilian population, described below.

Counter-adjustment to civilian population to account for change in HAF numbers

Every year there is a net transfer of people between the civilian population and HAF: this can be either positive or negative, depending on whether the size of HAF gets larger or smaller.

The total size of HAF is derived each year from data provided by Defence Analytical Service and Advice (DASA). However, any change to the size of HAF needs to be offset by a counter-adjustment to the civilian population.

The counter-adjustments at national level need to be broken down to LA level as ONS does not have specific data on the origin or destination LAs of people joining or leaving HAF. In addition the moves will not be picked up in the standard internal migration estimates as these are based on NHS GP registrations, whereas large HAF bases have their own medical system.

For the revised LA-level back series the revised national counter-adjustments were distributed across LAs in proportion to the distribution of the HAF population. This means that new personnel are taken from areas with a HAF presence, and departing personnel are placed back into the civilian population in HAF areas. It is recognised that the reality is more complex, but this approach is considered superior to distributing the counter-adjustments evenly across the country. Moreover, there are insufficient data available to create a 'real world' distribution.

Service dependents overseas (SDOs)

Every year there is a change in the number of SDOs – specifically dependants of HAF personnel who are accompanying the personnel overseas. Each year's change to the number of SDOs is offset by a counter-adjustment to the civilian population of England and Wales.

The revised national back series had no change to the number of SDOs. However, the previous back series distributed SDOs to LA level based on the distribution of the total population. As with the adjustment to take account of change in HAF numbers, the revised back series has distributed the SDOs using the distribution of HAF. This is logical: if SDOs accompany the service personnel overseas, it would seem likely they would remain with them upon return to England and Wales.

ii) Foreign Armed Forces (FAF)

The revised national back series had no changes to the number of FAF personnel present in England and Wales. However, recent analysis of 2011 Census data revealed that the FAF total for Harrogate LA in the mid-2011 estimates was about 400 too low, although the figures were satisfactory for all other LAs.

The difference in Harrogate has been incorporated into the revisions step by step from the mid-2009 estimates onwards so the Harrogate FAF total does not have an unusually large step change in any one year. However, in each set of estimates the subnational FAF figures have been constrained to the published national back series totals. To achieve this the FAF totals for other LAs have also been scaled down slightly (by about 5% each in the mid-2011 estimates, less in mid-2010 and mid-2009).

There has been no impact on the overall mid-2011 population of each LA. The components of change for the mid-2011 estimates had not previously been published and therefore the FAF members excluded due to the scaling of the FAF component have instead been included as part of the civilian population, and will be switched back in the processing for the mid-2012 estimates.

There has been an impact on LA population sizes in the revised mid-2009 and mid-2010 estimates. Harrogate has an increased population (reflecting the presence of the additional FAF personnel) and other LAs with FAF have had their populations slightly reduced.

8. School boarders

The former mid-year estimates included a specific component for school boarders. Data were obtained on the number of pupils in each boarding school and these were added to the population of the relevant LAs. A counter-adjustment was applied to represent the fact that the boarders were no longer in their home LAs, and this was distributed across the country.

The assumption behind this was that boarders would usually remain registered with a NHS GP in their home LA, rather than transferring to one in their school LA, and that therefore they would not be picked up in the usual internal migration data. However, recent research has shown that this assumption is no longer applicable: in practice boarders usually do register with a GP in their school LA, and this is Government policy.

The consequence of this is that people either starting or finishing at a boarding school in a different LA have their move double-counted, causing error. To overcome this, the revised mid-2002 to mid-2010 estimates do not have a specific school boarders component.

The effect of removing the double-counting is generally small: for example, in the rolled-forward mid-2011 estimates no LA was affected by more than 500 people.

9. Cross-border migration correction

The revised subnational back series contains small corrections to bring the LA-level estimates of cross-border migration into line with ONS's official estimates over the decade. This is the same approach as in the national revisions, where there was a net reduction of 2,400 in the flow from England and Wales to Scotland and Northern Ireland over the decade.

10. Removal of historic processing adjustments

The revised national back series added a total of 800 to the estimates for the years ending mid-2002 to mid-2006 due to the removal of small historic processing adjustments which had not been allocated to any specific cause. This very minor difference has been distributed across the country at LA level.

11. Other

General

Once all the components of difference identified above have been added to the rolled-forward mid-2011 estimate, for each age and sex group within each LA the remaining difference from the Census-based mid-2011 estimates has been classed as 'Other'.

As at national level, the Other component has been split across the decade on a cohort basis, meaning that it takes account of the fact that individuals age as the decade progresses. An example of how this works is as follows:

1. Suppose that in Borsetshire, once all identified components have been added, the rolled-forward mid-2011 estimates still have 100 fewer females aged 28 than do the Census-based mid-2011 estimates. Therefore the Other component for 28-year-old females is 100.
2. It is assumed that of this 100, a total difference of 10 has arisen in that cohort over each year of the decade.
3. Therefore, although the mid-2011 total is not changed, it is assumed that 10 of the difference arose in the year-ending mid-2011, among females who were aged 28 in mid-2011.
4. 10 arose in the year ending mid-2010 among those females aged 27 in mid-2010.
5. 10 arose in the year ending mid-2009 among those females aged 26 in mid-2009 etc.

Where an Other component is not a multiple of 10, it is split as evenly as possible across the decade. For those cohorts aged under 10 in mid-2011 the remaining difference has been allocated evenly to the years in which those children were alive.

Factors contributing to the Other component

The Other component in each LA is likely to be due to a combination of potential inaccuracies in any of the following:

- Internal migration. Apart from the changes to school boarder methods and the small revision to cross-border flows described above, migration between LAs in the UK has not been changed in the back series. However, in practice some moves (for example, those of young people finishing Further Education courses) are difficult to estimate accurately.
- International migration. There are two aspects here:
    - Any of the figures used for the revisions has scope for inaccuracy. This applies both to figures that have been revised and those that are unchanged, and could affect either immigration or emigration.
    - The new methods for distributing immigration down to LA level have only been applied to the years ending mid-2006 onwards. Had it been possible to apply them to the earlier years in the decade this would have led to different figures for each LA.
- The mid-2001 population estimates, which are the starting point for the series of estimates over the decade.
- The 2011 Census estimates. 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 potential for some uncertainty around the estimate. There is also the potential for other unidentified biases to have occurred.
- Prisoner definitions. The mid-year estimates up to mid-2010 define a ‘prisoner’ as somebody 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 tend to increase the population of LAs with prisons, potentially by several hundred, and slightly reduce it in other LAs.

- Any other component of population estimates over the decade.

Comparison with the national Other component

At national level the Other component accounted for 134,500 of the difference between the rolled-forward and Census-based mid-2011 estimates.

The Other component for each LA is independent of the national Other total. This is because the causative factors will impact each LA differently, and there are also factors (such as internal migration) which do not impact on the national total at all. Similarly, within each LA each age/sex class may have Other components of very different sizes, with the Other components for young adult ages (those ages where people migrate most) being likely to be highest.

In the majority of cases the size of the Other component is less than that of the difference between the rolled-forward and Census-based mid-2011 estimates, but in some cases it is larger. Cases where it is larger do not necessarily mean that the identified components of difference are incorrect; rather it may simply mean that there is a remaining larger inaccuracy elsewhere – for example, in internal migration.

Over the country as a whole, the Other totals at LA level sum to the national total. This is because all the factors which only impact at subnational level have offset each other across the country, leaving the national total.

12. Oadby and Wigston adjustment

Quality assurance revealed that applying the methods described above would result in an incorrect age distribution for Oadby and Wigston LA, resulting from complexities in estimating internal migration of students. The result was a tendency for substantial underestimation of the number of teenagers below student age, and an overestimate of people in their early 20s. These issues affected both sexes, but were more notable for males than females.

The issues were resolved by a special adjustment. This led to much-improved male and female age distributions for Oadby and Wigston, and also increased the number of people in their 20s in neighbouring Leicester, which is where most students in Oadby and Wigston move to after their first year of study. There were also very minor (negligible) adjustments in other LAs across the country.

Full details of this issue can be provided on request. However, it should be noted that only the mid-2002 to mid-2010 estimates are affected, and the mid-2011 estimates for Oadby and Wigston, Leicester and all other LAs are unchanged.

13. Case studies

The following case studies consider how successfully the attributable components of difference (all those apart from Other) have explained the gap between the rolled-forward and Census-based mid-2011 estimates in Newham, Luton and Manchester. In each case both the LA-level gap and the gap for individual age/sex groups are considered.

When interpreting the age/sex graphs in each case study, it is important to remember that any difference that has not been explained by the attributable components will instead have been filled across the decade by the Other component, so ensuring that the official mid-2011 total has been reached.

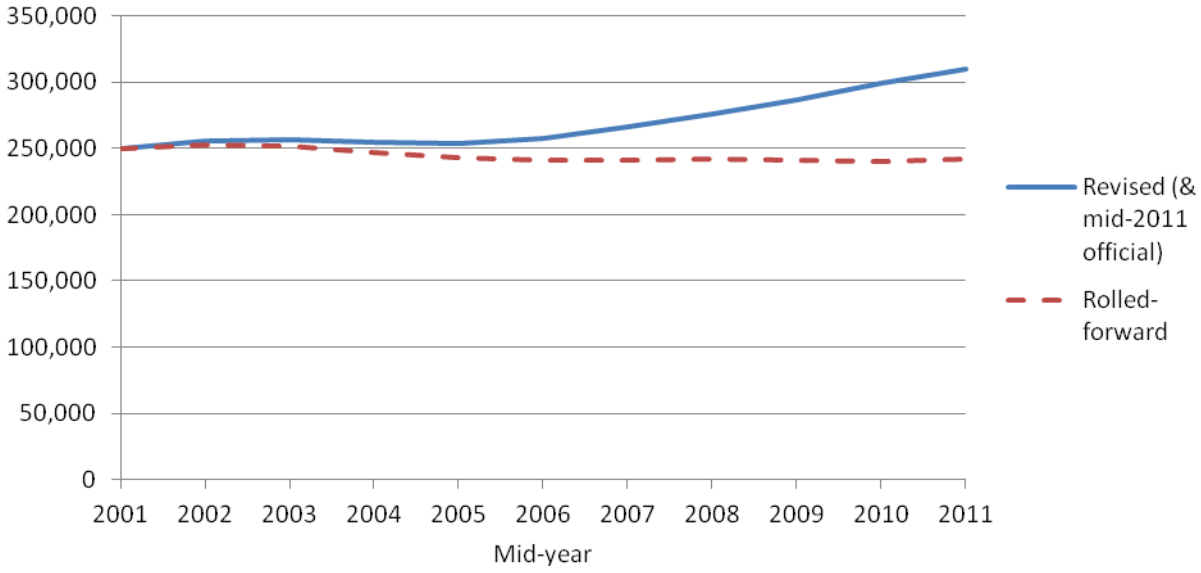
13.1 Newham

Newham is an urban LA in the north-eastern part of Inner London, and was one of the host boroughs for the 2012 Olympics and Paralympics. In mid-2001 its population was 249,400, but by the Census-based mid-2011 estimates it had risen by 61,000 to 310,500. This rise of 24% was the second highest percentage increase in the country, surpassed only by a 27% rise in the adjacent LA of Tower Hamlets.

However, the rolled-forward mid-2011 estimates gave a very different picture. They suggested that Newham's population had declined to 242,400 – a fall of 3%. This difference of 68,100 between the rolled-forward and Census-based estimates was the largest in the country.

Figure 1 shows how the difference has been added to the rolled-forward estimates across the decade in order to create the revised series.

Figure 1: Rolled-forward population estimates for Newham, mid-2001 to mid-2011, compared with the revised mid-2002 to mid-2010 estimates and the official (Census-based) mid-2011 estimates



This shows that most of the difference has occurred in the second half of the decade. The reason for this is demonstrated by the components of difference (Table 5).

Table 5: Components of difference for Newham, mid-2002 to mid-2011 estimates (totals)

Factor	Impact on difference	Remainder
Initial difference	n/a	-68,100
Switch to new method immigration estimates for years ending mid-2006 to mid-2011	41,400	-26,600
Consequent impact on emigration	-2,000	-28,600
EU8 immigration adjustment	6,000	-22,600
Republic of Ireland migration roll-back	500	-22,100
Sum of remaining attributable components	200	-21,900
Other	21,900	0

Note: totals may not sum due to rounding.

Table 5 shows that although the EU8 immigration adjustment makes an important difference, the predominant attributable component of difference for Newham is the introduction of the new immigration methods for the years ending mid-2006 to mid-2011. This is the main reason why the revision is concentrated in the second half of the decade.

Newham's Other component of 21,900 – the unexplained part of the difference – is also large, but one contributory factor could be that new method immigration estimates were not available for the years ending mid-2002 to mid-2005. The new methods added extra immigrants to Newham for all years in which they were applied, so if they could also have been applied to the earlier years then the Other component could potentially have seen substantial reduction. However, it is also possible that any combination of the other possible 'Other' factors (described in Section 11) could have had a large effect.

Another way of considering the impact of the attributable components of difference is to look at Newham's population by age and sex. (Figures 2 and 3)

Figure 2: Age profile of males in Newham, mid-2011

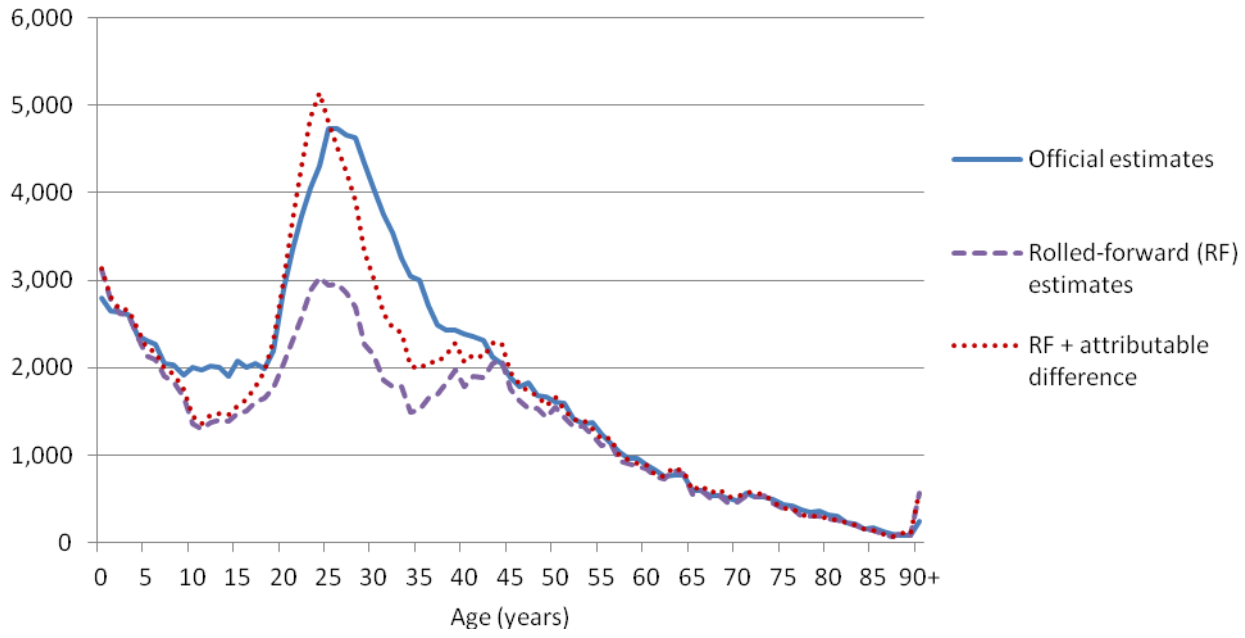
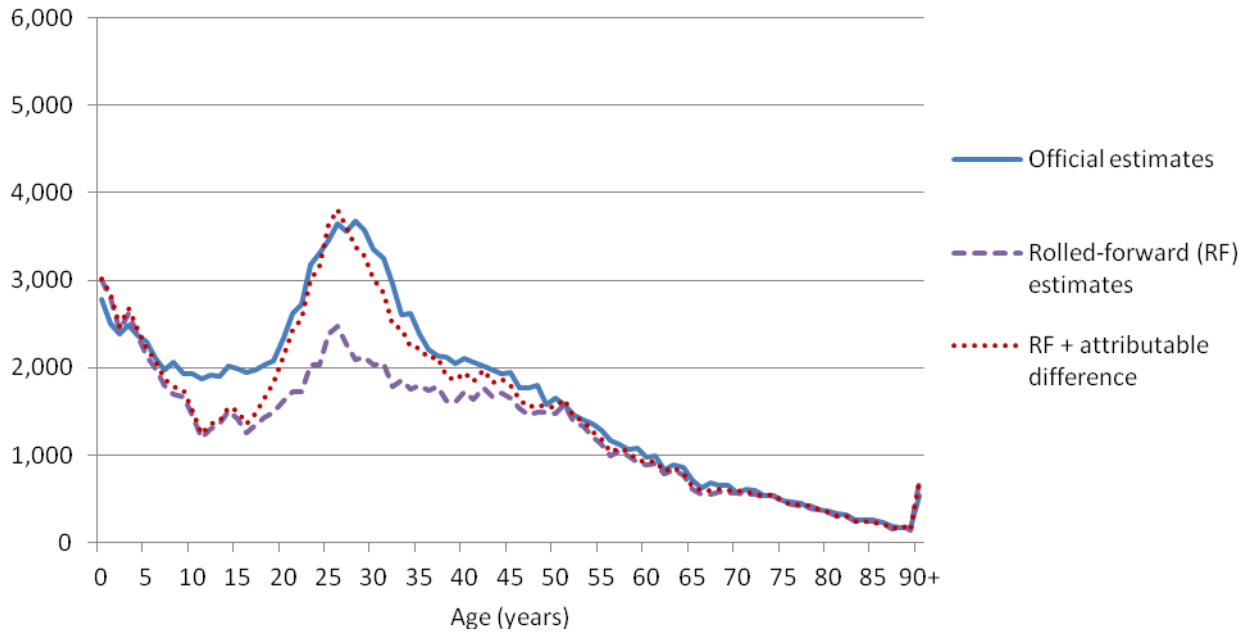


Figure 3: Age profile of females in Newham, mid-2011



In both figures the official (Census-based) mid-2011 estimates are higher than the rolled-forward estimates for almost all ages, and the addition of the attributable components of change closes the gap.

For both sexes the large underestimate of people in their 20s has disappeared, and especially for men in their mid-20s there is now a slight overestimate (meaning a negative Other component is needed to bring the numbers back to the official totals).

For people in their 30s the underestimate has been reduced, especially for females. However, for both sexes there has not been much reduction in the underestimate of people in their teens, meaning the Other component for those age groups is still large.

The methods paper for the national revisions referred to 2011 Census data that suggested that there may be inaccuracy in the methods for estimating the age distribution of international immigrants; this applies to both the former estimates as well as the revised estimates used in the mid-2002 to mid-2010 revisions.

Specifically, the Census data suggested that too many people might be being allocated to the young adult age categories, and too few to the child and slightly older adult age groups. If this is indeed the case it could be an important contributor to the Other component in Newham. Another factor is that the 2001 Census, informing the mid-2001 base for the rolled-forward series, may have underestimated the number of children then aged under 10.

Newham is also an area with a higher than average concentration of students. Such areas are liable to have greater error in internal migration estimates due to the fact that young people, especially males, are liable to take much longer to register with a new GP when they change LA. This lag time, plus the ageing of the affected cohorts through the decade, may have an impact on cohort sizes into the 30s age groups.

The new methods for distributing immigration to LA level are to become the standard for future mid-year estimates, starting with the mid-2012 estimates. ONS is also intending to introduce improved methods for estimating the age and sex distribution of immigrants, and for estimating moves of students after they finish their studies. Therefore it is expected that

immigration estimates for Newham will be much more accurate over the next decade, substantially improving the overall population estimates.

13.2 Luton

Luton is an urban LA comprising the industrial town of Luton in the East of England region. In mid-2001 its population was 185,900, but by the Census-based mid-2011 estimates it had risen by 17,800 to 203,600.

In the rolled-forward mid-2011 estimates the population was 204,000, which was very similar. The components of difference offset each other as follows, meaning that the Other component was just -100.

Table 6: Components of difference for Luton, mid-2002 to mid-2011 estimates (totals)

Factor	Impact on difference	Remainder
Initial difference	n/a	400
Switch to new method immigration estimates for years ending mid-2006 to mid-2011	-3,800	-3,400
Consequent impact on emigration	300	-3,100
EU8 immigration adjustment	2,500	-700
Republic of Ireland migration roll-back	500	-100
Sum of remaining attributable components	300	100
Other	-100	0

Note: totals may not sum due to rounding.

At an LA level the attributable components appear to have explained the difference. However, this does not exclude the possibility that there are other inaccuracies in specific components of the estimates, just that they also happen to have offset each other.

Figures 4 and 5 show the impact of the components of difference by age and sex:

Figure 4: Age profile of males in Luton, mid-2011

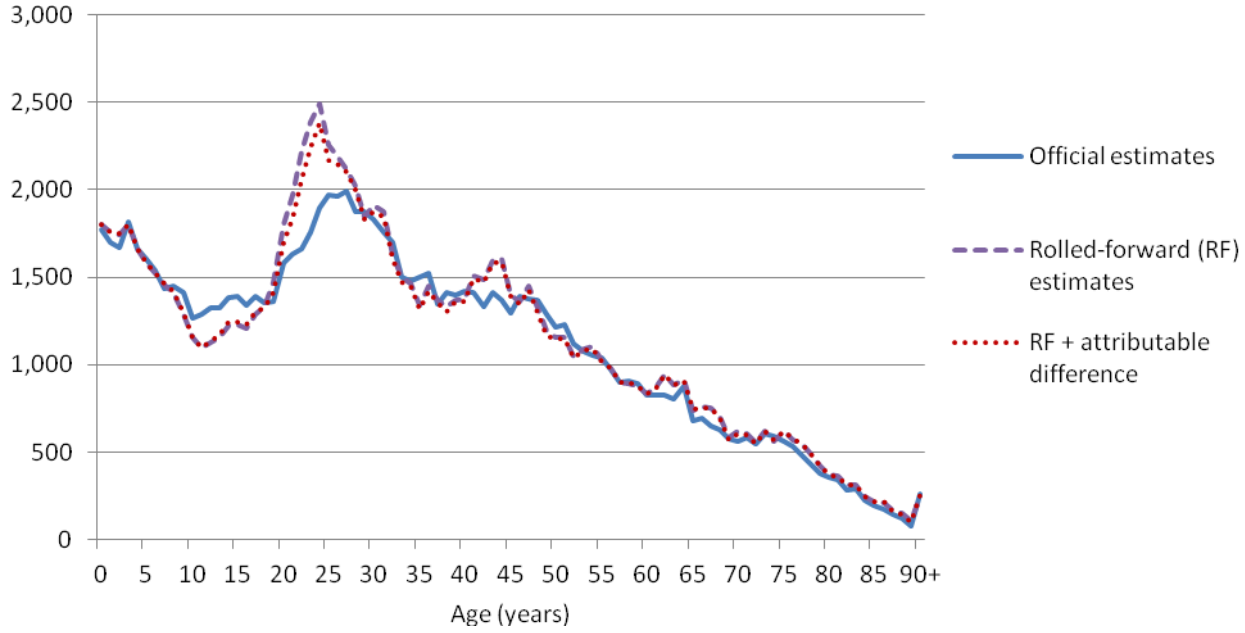
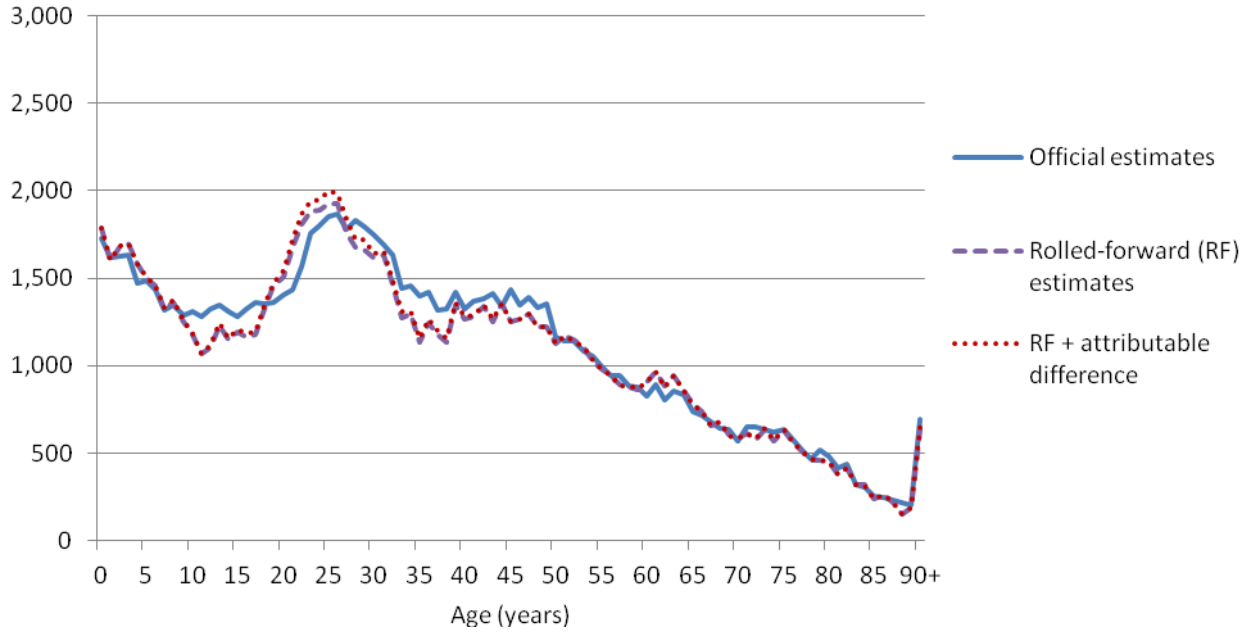


Figure 5: Age profile of females in Luton, mid-2011



In both figures it can be seen that the attributable difference has made very little change to individual age/sex groups. As with Newham there is still an underestimate of the number of people in their teens. There is also a slight underestimate of females in their 30s. However, for both sexes there is an overestimate of the number of people in their 20s, especially their early 20s.

Possible explanations of this are the factors which may have had an impact in Newham: an incorrect age distribution for immigrants, a possible underestimate of young children in the 2001 Census, and any errors in internal migration, although it should be noted that Luton's student population is closer to the national average than that of Newham (which has a comparatively large student population), so any 'student effects' should be smaller.

13.3 Manchester

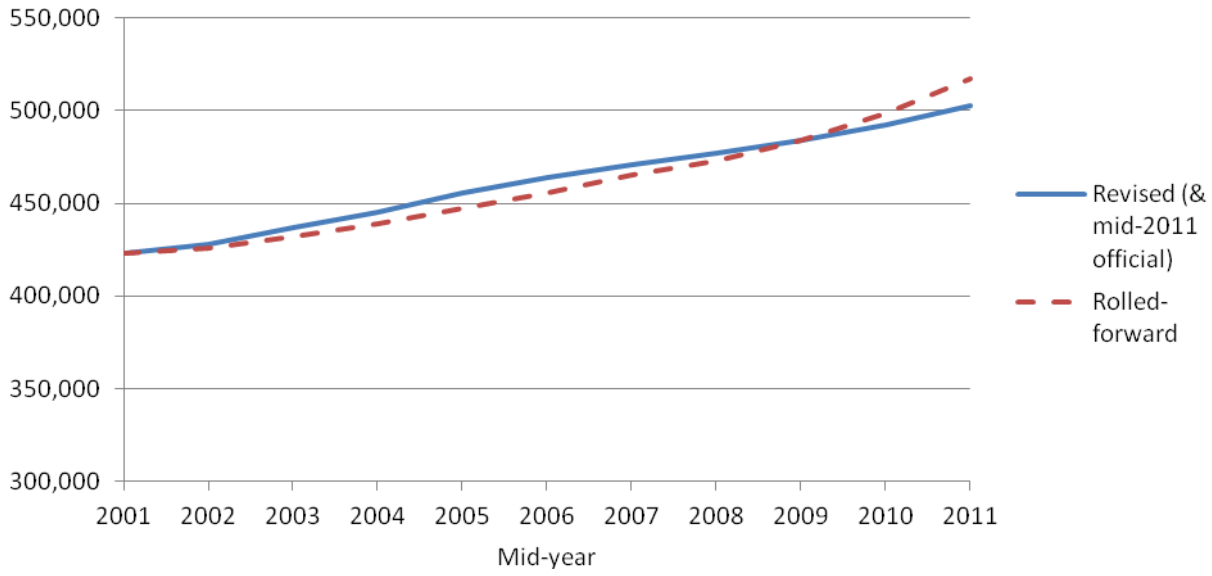
Manchester is an urban LA in the heart of the Greater Manchester conurbation in the North West region of England. It is a major regional centre with a large student population.

In mid-2001 its population was 422,900, but by the Census-based mid-2011 estimates it had risen by 80,000 to 502,900. This rise of 19% was the highest percentage increase in any LA outside of London.

However, the rolled-forward mid-2011 estimates suggested that Manchester's population was 517,000, an additional 14,100.

Figure 6 shows how the difference has been added to the rolled-forward estimates across the decade in order to create the revised series.

Figure 6: Rolled-forward population estimates for Manchester, mid-2001 to mid-2011, compared with the revised mid-2002 to mid-2010 estimates and the official (Census-based) mid-2011 estimates



Note: the y-axis has been started at 300,000 to emphasise the subtle difference between the lines.

This shows that there is little overall difference between the revised and rolled-forward series. During the early part of the decade the revisions slightly increase Manchester's population, but in the latter part of the decade the difference narrows, and by mid-2010 the revised estimates are clearly lower. This pattern can be explained by inspection of the components of difference (Table 7).

Table 7: Components of difference for Manchester, mid-2002 to mid-2011 estimates (totals)

Factor	Impact on difference	Remainder
Initial difference	n/a	14,100
Switch to new method immigration estimates for years ending mid-2006 to mid-2011	-40,000	-25,900
Consequent impact on emigration	1,900	-24,000
EU8 immigration adjustment	4,400	-19,600
Republic of Ireland migration roll-back	1,400	-18,200
Sum of remaining attributable components	-400	-18,600
Other	18,600	0

Note: totals may not sum due to rounding.

For the first part of the decade the main component of difference is Other, which increases the number of people in compared with the rolled-forward estimates. However, for the second half of the decade the switch to the new method immigration estimates becomes the dominant change, and this causes the revised estimates to grow more slowly than the rolled-forward estimates.

Overall the impact of the attributable components of difference is to turn an overestimate of 14,100 into an underestimate of 18,600. However, it should be noted that although the Other component of 18,600 is large in absolute terms (8th largest magnitude in any LA), Manchester is a very large LA meaning that it is much less notable as a percentage of the Census-based population (56th largest magnitude).

There are a number of factors which may be contributing to the Other component. It could be that the new methods have reduced immigration too far. However, it is also possible that the mid-2001 starting point of the rolled-forward series was too low. A third factor is that in an LA with a large number of students there has been more scope for inaccuracy in the internal migration estimates.

Figures 7 and 8 show the impact of the components of difference by age and sex. For both males and females, the attributable components of difference substantially reduce the overestimate for people in their 20s. However, as with Newham and Luton there is still a large underestimate for the teenage years, requiring a large Other component. Again, this may be due to a combination of a general underestimate of immigration of people in those cohorts, and a potential underestimate of under 10s in the 2001 Census, which would have been carried forward through the decade.

For women aged in their 30s and 40s the rolled-forward estimates are lower than the Census-based estimates, and the attributable components of difference make this issue slightly larger. This issue also occurs to a lesser extent for men in that age group. One possible factor could again be an incorrect age distribution of immigrants.

Figure 7: Age profile of males in Manchester, mid-2011

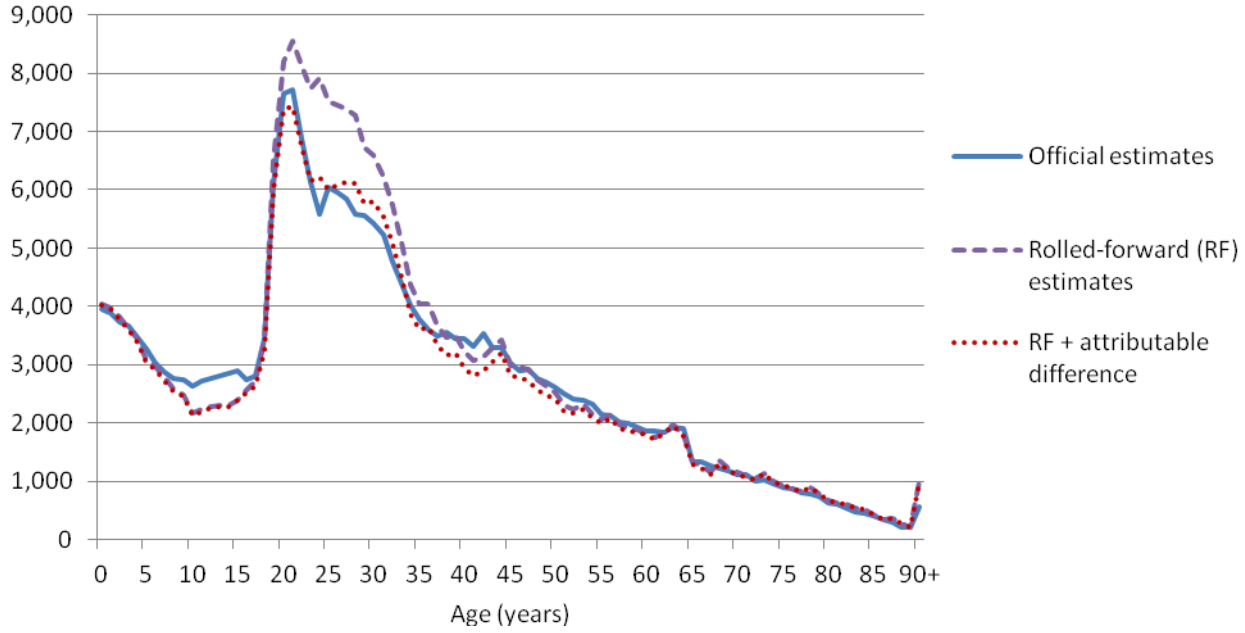
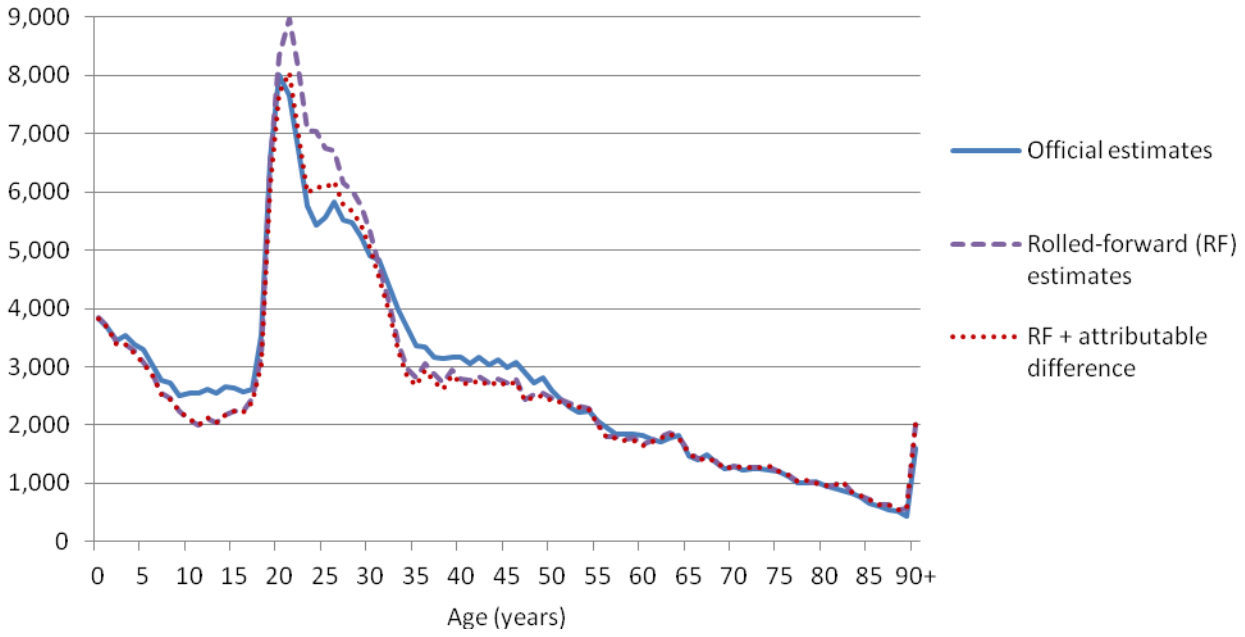


Figure 8: Age profile of females in Manchester, mid-2011



14. Conclusion and future work

The research informing the revised national and subnational estimates for mid-2002 to mid-2010 has sought to identify the cause of the difference between the mid-2011 population estimates rolled forward through the decade from 2001, and the official mid-2011 estimates based on the results of the 2011 Census.

At national level the main driver of the difference is believed to be underestimation of net international migration over the decade, with underestimation of the number of immigrants from the EU8 countries of central and eastern Europe being particularly important.

This underestimation of national net international migration has also had an important impact at LA level. However, for some LAs (including all three case studies) a more important issue has been the way in which international immigration has been broken down to LA level, with the application of new methods based on administrative sources making a substantial difference.

Taking all attributable components of difference into account, some LAs are left with a small Other (unattributable) component whereas others have a large Other component, sometimes larger than the original difference. This may be because of inaccuracy in the attributable components, most of which involve estimation and assumptions. However, it is also likely that other, unattributed factors have made a substantial contribution to the difference.

This publication of revised mid-2002 to mid-2010 subnational estimates offer a much improved and continuous series between the mid-2001 and official mid-2011 estimates, and ONS does not intend to make any further revisions.

However, work is continuing to develop the methods further:

- As indicated, the new method for estimating international immigration to LAs will be the standard.
- The mid-2012 estimates, due in June/July 2013, are also likely to feature a new method for estimating the age and sex distribution of immigrants, as well as refinements to the methods used to estimate moves of students after they finish their studies.

Other research is set to inform improvements in future years, ensuring that ONS's population estimates continue to be of the highest possible quality.

References and related publications

Data:

ONS (September 2012): *Population Estimates for England and Wales, Mid-2011 (2011 Census-based)*, http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2011--2011-census-based-/index.html.

ONS (September 2012): *Actual and percentage differences between rolled-forward mid-2011 population estimates and 2011 Census-based MYEs*, http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru-/difference-between-rolled-forward-mid-2011-estimates-and-2011-census-myes.xls.

ONS (October 2012): *Local authority 2011 Mid Year Rolled Forward Estimates based on old and new method by single year of age and sex*, http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru-/rolled-forward-mid-2011-population-estimates--old-and-new-methods-.xls.

ONS (December 2012): *Population Estimates for England and Wales – Mid-2002 to Mid-2010 Revised (National),* http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2002-to-mid-2010-revised--national-/index.html.

ONS (April 2013): *Population Estimates for England and Wales – Mid-2002 to Mid-2010 Revised (Subnational)*, http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2002-to-mid-2010-revised--subnational-/index.html.

ONS (April 2013): *Components of difference underlying the revised mid-2002 to mid-2010 population estimates,* http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru-/components-of-difference-for-revised-mid-2002-to-mid-2010-population-estimates.xls.

Papers:

ONS (2007): *Population Statistics: A Revisions Policy*, http://www.ons.gov.uk/ons/guide-method/revisions/revisions-policies-by-theme/population/population-statistics-revision-policy.pdf.

ONS (September 2012): *Examining the difference between the rolled-forward mid-2011 population estimates and the 2011 Census-based MYEs at local authority level,* http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru-/examining-the-difference-between-the-rolled-forward-mid-2011-population-estimates-and-2011-census-based-myes.pdf. Note that the comparisons in the paper are based on rolled-forward estimates which already take into account the new method for allocating immigration to local authorities.

ONS (December 2012): *Methods used to revise the national population estimates for mid-2002 to mid-2010,* http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/population-statistics-research-unit--psru-/methods-used-to-revise-the-national-population-estimates-for-mid-2002-to-mid-2010.pdf.