Explaining the Difference between the 2011 Census Estimates and the Rolled-Forward Population Estimates

Executive summary

The England and Wales March 2011 rolled-forward population estimate is 476,000 lower than the 2011 Census estimates (see Table 1). The difference is made up of 144,000 fewer males and 332,000 fewer females and in total is 0.9 per cent of the England and Wales 2011 Census estimate of 56 million people.

Table 1
The difference between census day and the revised rolled-forward population estimates

<table>
<thead>
<tr>
<th></th>
<th>Revised rolled-forward population estimates</th>
<th>2011 Census estimates</th>
<th>Difference to explain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>55,600</td>
<td>56,076</td>
<td>476</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td>27,429</td>
<td>27,573</td>
<td>144</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td>28,171</td>
<td>28,503</td>
<td>332</td>
</tr>
</tbody>
</table>

It is normal to find a difference between census estimates and population estimates. The high response rate for the 2011 Census combined with improved methods of estimating the population developed as part of the Migration Statistics Improvement Programme should ensure that this difference is kept to a minimum in the future.

In general terms, the population estimates over the 10-year period since 2001 Census have been prepared by starting with the 2001 Census-based population estimates, and every year adding on births, subtracting deaths, and adding in any net international migration that has taken place since 2001.

* New net international migration and births and deaths data became available which led to an improved estimation of the rolled-forward population. The improved estimation is 52,000 larger than the rolled forward estimate used for quality assuring the 2011 Census.
Births and deaths data are very reliable as they are underpinned by a comprehensive registration system. This means the difference of 476,000 between the rolled-forward population estimates and the 2011 Census estimates must be due to the 2001 Census underestimating the population and net international migration being underestimated over the 10-year period. This report focuses on these two reasons.

Initial research suggests that 209,000 (45 per cent) of the difference might be attributed to an assumed shortfall in the 2001 census day estimates which has been carried forward in population estimates through the decade.

The remaining 267,000 (55 per cent) would then be attributed to an underestimation of net international migration over the decade.

Early investigations into differences at the local authority level have identified some issues with internal migration and with the population estimates of local authorities with special populations such as armed forces.

More detailed information on differences for England and Wales and at local authority level will be published in three other reports over the coming months:

1. September 2012 – a high level report at local authority level, in conjunction with the release of the mid-2011 population estimates
2. December 2012 – a detailed report for England and Wales
3. March 2013 – a detailed report at local authority level
1. Purpose of this report

The purpose of this report is to identify the differences between the March 2011 Census estimates and the March 2011 rolled-forward population estimates (referred to as the rolled-forward estimates for the remainder of this report) for England and Wales, and to suggest reasons for those differences. Some early findings at local authority level are also reported.

The Office for National Statistics (ONS) produces annual estimates of the resident population as at 30 June each year. These mid-year population estimates (MYEs) tell us how many people live in England and Wales as a whole and in each local area, and include information on age-sex structure.

The most authoritative population estimates come from the census, which takes place every 10 years in the UK, the most recent being on 27 March 2011. The previous census took place on 29 April 2001. Population estimates from the 2001 Census have been updated each year to produce MYEs for the years in between censuses.

This report addresses the following questions for England and Wales:

- how accurate have the MYEs been over the past decade?
- how might differences from the 2011 Census estimates be explained?

As outlined in the executive summary, more detailed information on differences for England and Wales and at local authority level will be published in three other reports over the coming months.

These publications will not only inform users as to the nature of and reasons for the differences, they will also inform ONS’s work to further improve the quality of future population estimates.

2. Population estimates used in the report

This report uses population estimates rolled forward to census day (27 March 2011), taking account of actual births and deaths, and international migration data up to that date. Due to the use of more recent data, these population estimates are 52,000 higher than the rolled forward-estimates used for census quality assurance. The estimates used for quality assurance were produced by extrapolating the growth rates during the year to mid-2010 up to census day. See Appendix A.

3. Total difference

The 2011 Census estimates are 476,000 higher than the rolled-forward estimates, which splits down into 144,000 males and 332,000 females (see Table 3.1). To put this difference into perspective, the census estimates are 0.9 per cent higher than the rolled-forward estimates overall, 0.5 per cent higher for males and 1.2 per cent higher for females.

The four components that can contribute to this difference are births, deaths, net migration and the 2001 Census population base. As the registration of births and deaths that occur in the UK is compulsory, ONS is confident in the accuracy of these data. There is no similar
administrative system for registering migration; this is estimated using survey data, which has limitations. Therefore, at the England and Wales level, the discrepancy of 476,000 is due to issues with quantifying net migration (primarily net international migration) accurately, and the 2001 Census base.

### Table 3.1
Increase in population between mid-2001 and 2011 Census: comparison of Census estimates and rolled-forward (Census Day) estimates

<table>
<thead>
<tr>
<th></th>
<th>Mid-2001 MYEs</th>
<th>Rolled-forward population estimates</th>
<th>Increase over decade</th>
<th>2011 Census estimates</th>
<th>Increase since mid-2001</th>
<th>Difference to explain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>52,360</td>
<td>55,600</td>
<td>3,240</td>
<td>56,076</td>
<td>3,716</td>
<td>476</td>
</tr>
<tr>
<td>Male</td>
<td>25,575</td>
<td>27,429</td>
<td>1,854</td>
<td>27,573</td>
<td>1,998</td>
<td>144</td>
</tr>
<tr>
<td>Female</td>
<td>26,785</td>
<td>28,171</td>
<td>1,386</td>
<td>28,503</td>
<td>1,718</td>
<td>332</td>
</tr>
</tbody>
</table>

There are two main factors considered to have contributed to the difference of 476,000:

- 267,000 (55 per cent) is due to potential inaccuracies in the measurement of net international migration over the decade
- 209,000 (45 per cent) is attributable to potential problems with the 2001 population base (the starting point for the current series of MYEs), which is considered to be too low due to the undercounting of certain age groups in the 2001 Census

### 3.1 2001 Census

The 2001 Census was the first to be directly adjusted for non-response. However, further work following the census to address issues relating to ‘missed men’ and other specific local authority issues meant that changes were made to mid-year estimates for particular local authorities. This led to population being added to the MYEs. Even after these additions, there remained an unexplained population change (UPC) of 209,000 people between the census base with these additions, and the rolled-forward MYEs from the previous census.

The 2011 Census provides further evidence on the quality of the 2001 Census, and suggests whether these unexplained differences relate to quality issues in the MYEs or with the 2001 Census base. For example, the number of 0 to 9 year-olds was lower in the 2001 Census than in the MYEs. Given the greater number of 10 to 19 year-olds estimated in the 2011 Census than expected from the rolled-forward MYEs, it is likely that some of that difference is explained by the 2001 Census base being slightly low for 0 to 9 year olds. Further details of potential inaccuracies in the population base rolled-forward from the 2001 Census are provided in Appendix B.
4. Comparison between the 2011 Census estimates and rolled-forward estimates by age and sex for England and Wales

This section is based on analysis of the differences between the 2011 Census estimates and rolled-forward estimates by age and sex, while accounting for the UPC of 209,000 carried forward from 2001.

4.1 Breakdown of the difference by five-year age group

As explained above, the overall difference between the rolled-forward estimates and the 2011 Census estimates is 476,000, comprising 144,000 males and 332,000 females.

Figure 4.1.1 explains the difference in terms of five-year age groups; a positive difference indicates that the 2011 Census estimates are higher than the rolled-forward estimates. The difference is far more pronounced for some age groups than others.

Figure 4.1.1
Difference between 2011 Census estimates and rolled-forward estimates by sex and five-year age group (age at 2011 Census)

This chart shows which five-year age groups have contributed the most to the difference between the 2011 Census estimates and the rolled-forward estimates.

The most substantial positive differences for males are in the 10 to 19 and 30 to 39 age ranges. The Census has more males in those age groups than the rolled-forward estimates. However, the opposite is true for males aged 20 to 29 where the census estimates are substantially below the rolled-forward estimates.

For females, the pattern is slightly different. The census has considerably more females at ages 10 to 19 and 30 to 39. However, there are also more females at ages 20 to 29.

For other ages (0 to 9 and ages above 40) the census estimates and the rolled-forward estimates are relatively close for both males and females.
4.2 Age/sex breakdown of the unexplained population change (UPC) in the 2001 Census

This report has already referred to the UPC of 209,000 following the 2001 Census (see Appendix B for further information). The age/sex distribution of the UPC, aged on 10 years to the 2011 Census, is indicated in Figure 4.2.1.

In this chart, where the lines are above the x-axis the UPC is causing an overestimate in the rolled-forward estimates (i.e., the 2001 rolled-forward MYEs were higher than the 2001 Census-based MYEs). Where they are below the x-axis the UPC is causing an underestimation.

Overall the impact of the UPC is that the MYEs have underestimated males by 238,000 and overestimated females by 29,000. The largest differences are due to substantially fewer males, especially at ages 40 to 49, and considerably more females aged 30 to 44.

Note that the ages 0 to 9 are not included because these people were not yet born at the time of the 2001 Census.

Figure 4.2.1
Unexplained population change resulting from 2001 Census (age at 2011 Census)
5. How much of the difference is due to net international migration?

There is a 267,000 difference which needs to be explained that seems to be attributable to net migration. This difference is also variable by age, as shown in Figure 5.1. There are additional males in the 10 to 19 and 30 to 39 age groups, and fewer males in the 20 to 29 and 40 to 49 age groups, while there are more females in the 10 to 19 and 25 to 44 age groups.

The results provide an indication of the extent in each age group to which net migration has been either underestimated or overestimated in the rolled-forward estimates. The fact that the census estimates are higher suggests that we have underestimated net migration of males and females currently aged 10 to 19. For women we have underestimated net migration in the 25 to 44 age groups. The fact that the census estimates are lower suggests that we have overestimated net migration of males currently in their 20s and 40s.

Overall, the difference of 267,000 reflects an overestimation of net international migration of men by 94,000 and an underestimation of net international migration of women by 361,000 over the decade.

Figure 5.1
Difference between census estimates and rolled-forward estimates due to other factors
6. Potential reasons for long-term international migration underestimation

There is a 267,000 difference that seems attributable to net migration. The method to calculate long-term international migration (LTIM) for England and Wales is summarised by the equation:

\[ \text{LTIM} = (\text{International Passenger Survey (IPS)} + \text{visitor switchers} + \text{asylum seekers}) - \text{migrant switchers} \]

A long-term international migrant is defined as ‘someone who changes their country of usual residence for at least 12 months’. For further information about the LTIM methodology and definitions, and for net migration patterns over the decade, see Appendix C.

Substantial improvements have been made to the LTIM methodology in order to adjust for changes in passenger travel behaviour over the decade and to meet ONS’s commitment to continually improve data quality. At England and Wales level, changes to the method in the middle to latter part of the decade resulted in improvements to the:

- measurement of migrant and visitor switcher flows
- measurement of Republic of Ireland (ROI) and Northern Ireland flows
- IPS survey design, sampling and weighting process

As a result of these improvements, LTIM estimates have become more statistically robust in recent years. The difference between the 2011 Census estimates and the rolled-forward estimates is therefore most likely due to underestimation of LTIM prior to these improvements taking effect. This is broken down to:

- around 35,000 underestimation of migrant and visitor switcher flows
- around 55,000 underestimation of ROI flows
- around 175,000 due to underestimation from the IPS

6.1 Underestimation of visitor and migrant switchers

Visitor switchers are defined as those respondents who did not intend to stay in or leave the UK for longer than a year but subsequently did. In 2004, new IPS questions were introduced in order to establish more robust estimates for this group. The estimates were not backdated due to data availability. Assuming that migration patterns of these subgroups were similar in the early years of the decade, backdating the revisions would have resulted in an increase over the decade in net migration of an estimated 35,000.

Similarly, this improvement meant that migrant switcher (those respondents who intended to stay for longer than a year but subsequently did not) estimates were also more statistically robust. The impact of not backdating these improved estimates for net migration has been found to be negligible.

6.2 Underestimation due to Republic of Ireland flows method revisions

In 2008, improvements were made to the measurement of flows to and from the ROI. Again, due to data availability it was not possible to backdate these. Assuming that migration patterns of these subgroups were similar in the early years of the decade, then backdating
the revisions would have resulted in an increase over the decade in net migration of an estimated 55,000.

6.3 Underestimation due to the IPS

Method changes to the LTIM visitor and migrant switcher flows and Irish flows account for a probable 90,000 underestimation of migration over the decade (see Appendix C), but the largest contributor to the LTIM figures is the IPS. It seems that the remaining 175,000 is therefore attributable to IPS underestimation. This could be due to underestimating immigration or overestimating emigration. The emigration component is not examined in any detail in this current report, but will be discussed in a more detailed report for England and Wales due to be published in December 2012.

6.4 Comparison of long-term immigration estimates with other relevant data sources

Long-term immigration figures (broken down to IPS estimates where relevant) were compared with other datasets to identify whether there has been an IPS underestimation over the last decade.

England and Wales long-term international immigration figures were compared to ‘New Flag 4s’ by year extracted from the NHS Patient Register (PR). The PR documents new patients whose previous address was not in the UK by assigning a ‘Flag 4’ to their record. This Flag 4 indicator is removed when the patient moves and re-registers with a new address (and thus becomes an internal mover).

Figure 6.4.1 shows a comparison of the PR New Flag 4s and ONS’s published long-term immigration estimates for England and Wales from 2001 to 2010. From 2001 to 2005, the number of New Flag 4s is lower than the immigration estimates, but both follow a similar upward pattern. From 2005 onwards, the two cross and diverge. The New Flag 4s continue an upward trend to 630,000 in 2010 whereas the immigration estimates stay relatively stable over time, with 530,000 immigrants estimated in 2010. Over the decade, PR New Flag 4s totalled 5.3 million which is approximately 200,000 higher than ONS immigration estimates. This could indicate an underestimation of immigration by ONS, or it could be due to quality issues in the PR or, in fact, be due to some combination of both.

Data quality of the PR has improved over the decade but there are still some known issues. The PR is prone to variable levels of coverage across the UK and is thought to have a small national list inflation. List inflation is caused by people not being removed from the list when they leave the country. It should be noted that the PR may well include some short-term residents (those who come to the UK for less than 12 months) whereas these are excluded from the long-term immigration estimates. This may explain some of the observed difference.
Figure 6.4.1
Comparison of long-term immigration and PR New Flag 4 data for England and Wales (2001-2010)

It is difficult to quantify the exact scale of any differences between long-term international immigration and ‘New Flag 4s by year’ data given the definitional differences. However, the divergence between the two does occur in the middle of the decade, around the time when migration to England and Wales by citizens from central and eastern European countries increased.

The increase in migration of citizens from these countries was a result of the substantial European Union (EU) expansion that occurred in 2004. Eight countries from central and eastern Europe joined at the same time: Poland, Lithuania, Latvia, Slovenia, Slovakia, Czech Republic, Estonia and Hungary – these are termed the ‘EU8’ countries. The accession of these eight countries into the EU meant their citizens were then free to travel to, and work in, England and Wales. Given the financial circumstances of the EU8 countries at the time, it was expected that migration to England and Wales would increase for economic reasons.

6.4.1 EU8 accession and the impact on migration

The EU expansion did result in an increase in immigration from the EU8 population for work. Migrant worker figures in the IPS and the Lifetime Labour Market Database (L2) were examined. The L2 is a 1 per cent extract of data from the National Insurance and Pay As You Earn System (PAYE) and has over 750,000 individuals in the sample spanning over 30 years. For sample members, the L2 holds information on the date of registration; and for migrants, the self-reported date of arrival to the UK. The L2 also contains information on the number of weeks that a person has been active in a tax year.

The number of L2 migrant worker registrations from the EU10 countries (EU8 plus the additional 2007 EU2 accession countries, Romania and Bulgaria) was compared with IPS migrants who stated their reason for visit was to work. The L2 data include re-arrivals as well
as first time migrants. To limit definitional differences, those L2 records with more than a six-month difference between arrival and registration (as their main migration reason was less likely to have been to work) were excluded. Note that one disadvantage of the IPS is that it can only record someone’s stated main reason for immigrating, which may not correspond with their actual activity (for example, their main reason may be to join a partner, but they subsequently also get a job).

In 2004, the EU8 countries joined the EU and a sharp increase in the rate of L2 and IPS workers can be seen in this year (see Figure 6.4.2). However from 2005, the L2 (125,000) identifies substantially more EU8 workers than the IPS (50,000). Between 2005 and 2010 the L2 data identifies 300,000 more EU10 workers coming to England and Wales than the IPS. Even given some definitional differences, the evidence does suggest that the IPS did not measure the full scale of EU10 immigration in the middle of the decade.

**Figure 6.4.2**
England and Wales L2 and IPS EU10 migrant workers comparison (2001-2010)

Note: L2 figures are by tax year (up to March), IPS by mid-year (up to June)

6.5 IPS improvements

As discussed above, it is possible that the IPS did not adequately measure the rise in EU8 migration in the middle of the decade. Evidence presented here does suggest that IPS EU8 underestimation could explain the remaining difference of 175,000. While noting that the IPS may not have accurately measured this group in the past, substantial improvements have now been made both in terms of the IPS’s coverage and quality in the latter part of the decade.

In 2006, ONS set up an Inter-Departmental Task Force on Migration Statistics, and launched the Migration Statistics Improvement Programme (MSIP). The MSIP has been recognised as
a successful programme by demographic experts and users. A number of improvements were made to the quality of migration statistics, including improving the timing and accuracy of migration estimates for England and Wales and at local authority level.

Improvements to the IPS were implemented as a result of recommendations made by the Inter-Departmental Task Force on Migration Statistics. In 2008, the IPS was extended to cover regional airport travel to account for the increase in passenger traffic to and from these ports. Migration filter shifts were introduced to boost the number of migrant contacts at certain airports and additional shifts were introduced on sea routes and the Channel Tunnel Service. Other improvements were made to shift timings, as well as to the overall survey design and weighting process. These changes have resulted in the IPS data becoming more statistically robust, and, going forward have resulted in a substantial improvement in the quality of the international migration estimates.

6.6 Uncertainty in survey-based estimates

Between 2001 and 2010 net international migration based on IPS was estimated to be 1,508,000. However the uncertainty around migration net flows could be as great as ± 138,000 (using the 95 per cent confidence interval). This means in 19 out of 20 occasions, the true net flow total would most likely lie somewhere between 1,370,000 and 1,646,000 (had all migrants been surveyed). This uncertainty could therefore account for a substantial part of the difference.

6.7 Evidence of net migration underestimation by sex

The age and sex breakdown of the 267,000 difference due to migration (see Figure 5.1) suggests that LTIM has underestimated female immigration or overestimated female emigration. Alternatively, LTIM may have overestimated male immigration or underestimated male emigration. The migration component of the difference between the 2011 Census estimates and the rolled-forward estimates has been explained; however, evidence supporting a disproportionate sex bias in the LTIM methodology has not emerged. This issue will be considered in a subsequent report due for publication in December 2012, which will focus on England and Wales in more detail.

7. Identified census issues that impact on local authority population estimates

This section provides a high-level overview of the differences at local authority (LA) level between the 2011 Census estimates and the rolled-forward estimates.

Figure 7.1 summarises the difference between the March 2011 Census and the rolled-forward population estimates at LA level. For the majority of LAs, the difference between the 2011 Census estimate and the population estimate is relatively small with 89 per cent of LAs having a difference of less than 5 per cent and 91 per cent of LAs having a difference of less than 10,000.
The 20 LAs with the largest positive difference between the census estimates and the rolled-forward population estimates (that is, the census estimates are higher) are shown in Figure 7.2, together with the corresponding percentage difference. The bars represent the absolute difference in terms of the number of individuals. The crosses illustrate the percentage difference. For example, for Birmingham the numerical difference is approximately 31,000 people but as a percentage it is approximately 3 per cent of the population estimate.
The 20 LAs with the largest negative difference between the census estimates and the rolled-forward population estimates (that is, the census estimates are lower) are shown in Figure 7.3 together with the corresponding percentage difference.
It can be seen in Figures 7.2 and 7.3 that in some LAs the difference is large in absolute terms but is a small percentage of the overall population. For example, Leeds has an absolute negative difference of approximately 37,000 people, but a negative percentage difference of approximately 5 per cent. In other cases, the difference is quite large in both percentage and absolute terms.

These differences will be considered in more detail when comparisons can be made between the rebased mid-2011 population estimates and the rolled-forward mid-2011 population estimates later this year.
7.1 Possible reasons for the differences at local authority level

Some differences would be expected between the 2001 Census and the rolled-forward population estimates. To limit the number of possible reasons to be investigated, it is helpful to identify the five components of the methodology that may cause differences at the LA level. These are:

- the 2001 Census base
- births
- deaths
- internal migration, and
- international migration

Registrations of all births and deaths in England and Wales are collected by the General Register Office and are considered to be accurate. Therefore, any differences will be due to errors in either the 2001 Census base, international migration or internal migration.

The potential errors in the 2001 Census base and estimates of international migration identified at England and Wales level will also have had an impact at LA level, but this is not assessed in any more detail in this report. The following points summarise some of the internal migration issues that may contribute to the difference and, where developed, describe what steps are being taken to address the issues in the forthcoming population estimates based on the 2011 Census.

7.2 Examples of issues in the population methodology that may affect all LAs

7.2.1 Estimation of internal migration

Analysis of England and Wales population estimates has shown that in some areas there have been rises in sex ratios at young adult ages across the decade (ie an increasing proportion of men to women). While in some areas this has been driven by international migration, in many areas this has also been caused by internal migration (when people move between LAs within England and Wales), but the 2011 Census estimates do not show the same effects.

This suggests that there may be a problem with the estimation of internal migration for men within the mid-year estimates methodology. It is known that men are less likely to register with a GP and take longer to register when they move. This is illustrated by recent analysis of the Longitudinal Study (LS) which showed that when compared with women, men were less likely to have the same address recorded or be present on both the census and the National Health Service Central Register (NHSCR) in 2001⁷.

This known issue in GP registration behaviour may be affecting the estimation of internal migration for men and, consequently, the sex ratios at LA level of both the migration and population estimates. This therefore explains some of the differences at LA level between mid-year estimates and census. The effect is most noticeable in areas where there are more internal moves such as urban and university areas. It will not explain differences for the whole of England and Wales.

Figure 7.2.1 shows the number of internal moves by sex in the period between the 2001 and 2011 Censuses, and illustrates the possible underestimation of male moves. It shows that
there are fewer male internal moves when compared with female internal moves. This is potentially further evidence to support the theory that men are being underestimated in the internal migration estimates.

**Figure 7.2.1**
Internal migration moves across local authority boundaries within England and Wales during 2002 and 2010 based on GP registrations

Further research is needed to gain a better understanding of internal migration moves by sex and this will be investigated as part of the programme of improvement work planned for 2013. Further research is required to confirm if male GP registration patterns are creating a bias in the internal migration estimates.

**7.2.2 Estimation of student moves**
The comparisons between the 2011 Census and the rolled-forward estimates have shown that the rolled-forward estimates are too high for young people aged 25 to 29 and above in some university areas. This may be due to an underestimation of students moving out of university areas at the end of their studies which would lead to inflated estimates of adults in these age groups. The comparisons also indicate that the number of people in the population estimates aged 19 to 24 have been overestimated in some student areas (and possibly underestimated in other areas). Currently, an adjustment is made to the internal
migration estimates for student moves to and from university that are not captured using the internal migration methodology based on GP registrations.

The student adjustment component of the internal migration methodology will be reviewed in the light of the 2011 Census results.

7.3 Examples of issues in the population methodology that may affect some LAs

7.3.1 Estimation of armed forces

A definitional difference has been identified between the 2011 Census and the population estimates methodology, which affects comparisons between the census estimates and the rolled-forward estimates.

As a result of this definitional difference, census results for LAs containing armed forces bases are likely to show a decrease in the number of armed forces personnel, when compared with the rolled-forward population estimates. In addition, the enumeration of armed forces personnel at the permanent or family home will result in a slight increase in the census population across a large number of LAs. Refer to Appendix D for a proposed method to deal with this issue.

7.3.2 Estimation of school boarders

The population estimates methodology has assumed that most boarding school pupils remain registered with their home GP, meaning that their move to the LA where they attend boarding school is not captured. This issue is currently addressed by adding the population of each boarding school, provided by the Department for Education (DfE), to the population estimate of each LA. However, analysis has shown that GP registrations correlate strongly with DfE data at boarding school postcodes, suggesting that school boarders do usually register with a GP at their school address. This means that the addition of the school boarder component introduces duplication to the population estimates. Refer to Appendix D for a proposed method to deal with this issue.

7.3.3 Estimation of prisoners

There are definitional differences between the 2011 Census definition of a prisoner and the definition used for the population estimates – the 2011 Census includes prisoners on a sentence of six months or more irrespective of date of conviction, whereas the population estimates only include persons who have already served at least six months in prison. Recent research using data from the Ministry of Justice (MoJ) has found that the difference is around 20,000 people. It has been agreed that the 2011 Census-based and future population estimates will adopt the 2011 Census definition of a prisoner, with supporting data provided by the MoJ. Refer to Appendix D for a proposed method to deal with this definitional change.
8. Conclusion

The 2011 Census estimates for England and Wales are approximately 476,000 higher than the rolled-forward estimates. The main factors explaining this are believed to be error in the base population derived from the 2001 Census, and error in the estimation of international net migration over the intervening decade.

At LA level, a number of other factors are liable to contribute to over- or underestimation in the population estimates. These include methods for counting armed forces, school boarders, prisoners and, in particular, internal migrants.

Clearly any estimation process will inevitably lead to a degree of error, but by researching and understanding the differences that have arisen over the decade, ONS will be able to further improve its methods to ensure that population estimates in years to come are of the highest quality possible.
References


Appendix A: Improved method used to estimate the rolled-forward population estimates on census day

The original March 2011 rolled-forward estimate was produced by assuming that population change between mid-2009 and mid-2010 continued at the same rate for the period between mid-2010 and census day on 27 March 2011.

However, actual data that have more recently become available have indicated the rolled-forward estimates for census day should be higher than that originally calculated by 52,000. The main driver of this is higher international immigration than in the original rolled-forward (extrapolated) estimates.

Appendix B: Why the 2001 Census base may be inaccurate

The 2001 Census estimate was 1.1 million below the 2001 MYEs rolled forward from the 1991 Census. This discrepancy was accounted for in the following ways:

- 351,000 was explained by an estimated overcount in the 1991 Census base; this amount was subsequently removed from the 1991 base
- 305,000 was explained by inaccuracy in the estimation of international net migration between 1991 and 2001; the MYEs were subsequently reduced by this amount over the decade
- the 2001 base was increased by 193,000 due to a presumed 2001 Census underestimate of men aged 25 to 49. In total, 68 LAs were affected
- the 2001 base was increased by another 107,000 following the 'Local Authority Studies', whereby local authorities seen as having specific issues with their 2001 Census estimates were corrected
- in addition, the 2001 base was reduced by a further 25,000 as there was some overlap between the LAs initially adjusted using the ONS Longitudinal Study and the subsequent population studies adjustments. This left an unexplained population change (UPC) of 209,000 – that is, the 2001 base was 209,000 lower than expected for reasons that have not been explained
Table B.1
Explanation of the difference between the 2001 Census estimates and the 2001 rolled-forward MYEs

<table>
<thead>
<tr>
<th>Difference between 2001 Census estimates and rolled-forward MYEs</th>
<th>1,140,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991 adjustment</td>
<td>351,000</td>
</tr>
<tr>
<td>Migration adjustment</td>
<td>305,000</td>
</tr>
<tr>
<td>ONS Longitudinal Study adjustment</td>
<td>193,000</td>
</tr>
<tr>
<td>(and others in September 2003)</td>
<td></td>
</tr>
<tr>
<td>LA Population Studies</td>
<td>107,000</td>
</tr>
<tr>
<td>Longitudinal Study consequential adjustment (and other adjustments in September 2004)</td>
<td>-25,000</td>
</tr>
<tr>
<td>Remaining unexplained population change (UPC)</td>
<td>209,000</td>
</tr>
</tbody>
</table>

If the UPC figure of 209,000 is assumed to reflect people who were missed from the 2001 Census and subsequent adjustments – that is, the 2001 base should have been 209,000 higher, then it would account for around 45 per cent of the difference between the 2011 Census estimates and the rolled-forward estimates.

All of the adjustments leading to this UPC of 209,000 are inevitably subject to error, and it is possible that, if more information had been available at the time, additional adjustments might have been made. Note too that the 2001 Census result itself had a 95 per cent confidence interval of ± 104,000 due to sampling error – uncertainty arising from the need to estimate for people not included on a census questionnaire. Therefore, the 209,000 should only be regarded as a broad estimate of how much the 2001 base might be wrong.
Appendix C:

Net migration over the decade (2001-2010)

According to current ONS long-term international migration (LTIM) estimates, in the 10-year period between the calendar years 2001 and 2010, around 5 million people migrated to England and Wales. Immigration in 2001 was around 450,000. In 2004, immigration rose relatively sharply (from 480,000 to 550,000). It is likely that this was due to the EU expansion around that time. Since then, immigration has remained relatively stable.

Emigration ranged from 280,000 in 2001 and peaked in 2008 at 390,000. Around 3 million people migrated from England and Wales over the decade, notably fewer than those immigrating.

Table C.1
Published LTIM England and Wales figures 2001 - 2010

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immigration</td>
<td>449</td>
<td>485</td>
<td>480</td>
<td>549</td>
<td>523</td>
<td>549</td>
<td>524</td>
<td>528</td>
<td>507</td>
<td>532</td>
</tr>
<tr>
<td>Emigration</td>
<td>282</td>
<td>328</td>
<td>333</td>
<td>311</td>
<td>328</td>
<td>369</td>
<td>307</td>
<td>393</td>
<td>328</td>
<td>308</td>
</tr>
<tr>
<td>Net migration</td>
<td>167</td>
<td>158</td>
<td>147</td>
<td>238</td>
<td>195</td>
<td>180</td>
<td>217</td>
<td>135</td>
<td>179</td>
<td>224</td>
</tr>
</tbody>
</table>

How is long-term international migration measured for England and Wales?

The core component of the LTIM estimates is data collected via the International Passenger Survey (IPS). The IPS is an ONS sample survey of passengers who arrive at, and depart from, United Kingdom air and sea ports and the Channel Tunnel. IPS data are collected by face-to-face interviews with a random and voluntary sample of passengers as they enter or leave the UK.

IPS survey data are weighted up to be representative of total numbers travelling in and out of England and Wales. These figures are then used as the core component of LTIM estimates. Final LTIM estimates are produced once adjustments are made to account for asylum seekers, visitor and migrant switcher flows.

A strength of the IPS is that it is specifically designed to identify people who meet the UN definition of an international migrant, that is, someone who changes their country of usual residence for at least 12 months. This is also consistent with the ‘usual residence’ definition that is the basis for official population estimates. Producing outputs with definitional consistency is valuable as it gives users a common understanding of underlying concepts and definitions.8
Appendix D: Explanation of proposed methods to address internal migration issues in the population estimates

Armed forces

In 2001, armed forces personnel were instructed to be included on their spouse’s or partner’s questionnaire if they were usually resident at that address. Single members of the armed forces who were living on a base should have been enumerated at the base. However, the 2011 Census was designed so that members of the armed forces were enumerated at their ‘permanent or family home’, even if they were living at an army base for the majority of the time. This was a deliberate census design reflecting to a certain extent what actually happened in the 2001 Census but which is not consistent with the population estimates definition that armed forces should be counted in a place where they are usually resident.

As a result of this definitional difference, census estimates for LAs containing armed forces bases are likely to show a decrease in the number of armed forces, reducing their overall population when compared with the rolled-forward estimates. In addition, the enumeration of armed forces at the permanent or family home will result in a slight increase in the census estimate across a large number of LAs.

A proposed solution has been developed to address this definitional difference so that the rebased (census-based) population estimates for mid-2011 will use census data adjusted to include armed forces at their place of usual residence – an armed forces base if applicable, rather than their permanent or family home. This is in line with the European regulation for population data which states that members of the armed forces who live in a military barracks or camp should be estimated at the residence where they spend most of their daily period of rest. The adjustment will be made by applying a set of criteria to members of the home armed forces who have been enumerated at a permanent or family home but have also recorded a second address at an armed forces base.

This adjustment to the census data used in the rebased mid-2011 population estimates will mean that they will include armed forces using the same definition of residence as in previous MYEs.

School boarders

Data from the Department for Education have shown that school boarders are generally a small component of the population estimates with fewer than 6 per cent of 9 to 18 year-olds classified as boarders in 95 per cent of LAs. However, there is a small number of LAs in which boarding schools contribute up to 19 per cent of their population aged 9 to 18 years. Evidence suggests that these school boarders do usually register with a GP thereby introducing duplication to the population estimates. This duplication will have a relatively small impact overall but a larger effect in a small number of areas.

Following a review of the methodology, it is proposed that school boarders are to be removed as a separate component.

It is assumed that there will be no change in the number of school boarders between census day and mid-year and so this change in methodology will be implemented for the first time in
the mid-2012 population estimates. Therefore any error associated with the old methodology for school boarders will not affect the census estimates or the rebased mid-2011 estimates but will affect comparisons with the March 2011 rolled-forward estimates, as well as population estimates from previous years.

Prisoners

The population estimates currently count prisoners as people who have already served at least six months in prison. However, the 2011 Census defines prisoners as those on a sentence of six months or more irrespective of date of conviction.

Recent research using data from the Ministry of Justice (MoJ) has found that the difference between the 2011 Census definition of a prisoner and the definition used for the population estimates is around 20,000 people. To address this difference, it has been agreed that the 2011 Census-based and future population estimates will adopt the 2011 Census definition of a prisoner, with supporting data provided by the MoJ. This will effectively mean that both the census and subsequent population estimates will slightly increase the population in LAs with prisons, and decrease it in other LAs.

In addition, there are other more minor definitional differences between the census estimates of prisoners and the data available from MoJ. This could potentially mean that the census contains a few thousand ‘extra’ prisoners in all. For the rebased mid-2011 population estimates it has been proposed to redistribute these extra prisoners proportionally across the country by age and sex according to the non-prisoner population. This will not affect the figures for the whole of England and Wales, but will reduce counts slightly for LAs with prisons.

The existing series of estimates for prisoners used in the population estimates for mid-2002 to mid-2010 will be retained and so a step change will appear between the mid-2010 and mid-2011 population estimates in local authorities where prisons are present, following adoption of the census definition of a prisoner.