

Methodology Note on production of Super Output Area Population Estimates

October 2009

1. Introduction

The Office for National Statistics (ONS) produces estimates of the resident population of England and Wales. The most authoritative population estimates are produced every ten years and are based on the Census of Population and Housing. These are updated to produce mid-year population estimates in the intercensal period. The population estimates give a stock count of people living in local authority areas and health areas¹, and the composition of the population in these areas by age and sex. Further population statistics, including migration estimates, vital events (covering births, deaths, marriages and divorces) and population projections are also available.

In response to the increasing demand for small area statistics identified by initiatives such as New Deal for Communities, Best Value, the National Strategy for Neighbourhood Renewal and a review of customer needs for population statistics carried out by ONS, in April 2005 ONS published mid-2001 and mid-2002 Census Area Statistics (CAS) ward population estimates.

To reflect the new geographic hierarchy of Super Output Areas (SOAs) designed to improve the reporting of small area statistics in England and Wales, ONS has been publishing population estimates for Lower Layer Super Output Areas (LSOAs) and Middle Layer Super Output Areas (MSOAs). Estimates for mid-2001 to mid-2008 are currently available. ONS will for the foreseeable future be producing these estimates on an annual basis.

The mid-2001 SOA estimates largely reflect the 2001 Census counts but incorporate adjustments to reflect the 2001 local authority mid-year estimates. The methodology used to produce the SOA estimates for mid-2002 onwards differs to the method used to produce the local authority mid-year estimates, and is subject to further development. In view of this ongoing work the mid-2001 to mid-2008 SOA estimates are being published as 'experimental statistics'. Although these estimates are experimental and do have identified limitations, they do highlight the potential for new uses to be made of administrative data sources.

¹ Including Primary Care Organisations and Strategic Health Authorities in England, and Local Health Boards and Regional Offices in Wales

2. Methodological and data considerations

ONS has previously reviewed the methodology for producing small area population estimates in other National Statistics offices, and their availability within England and Wales from other providers. The availability and characteristics of potential data sources for use in an estimation model have been investigated, together with the identification of different demographic methods which could be used to produce small area population estimates. This led to the development of evaluation criteria for identifying the most suitable estimation model, comprising a method and data sources.

The different potential demographic methods identified for producing small area population estimates were evaluated against an identified set of evaluation criteria in order that a shortlist of methods could be identified to investigate further. The three methods which were shortlisted (Apportionment, Cohort Component and Ratio Change) are summarised in Appendix A.

An initial evaluation of these three methods was undertaken, and the Ratio Change method was identified as the preferred method. A further evaluation was undertaken in 2006, examining test estimates produced from each method over a three-year period. The Ratio Change method was again identified as the preferred method, however it was recognised that there wasn't overwhelming evidence for this method over the Cohort Component method, whilst the Apportionment method was the least favoured of the three shortlisted methods.

It is intended that a fuller further evaluation of these three methods will be undertaken when the results of the 2011 Census become available.

3. Population Definition

The population base from the 2001 Census underpins the mid-year population estimates resident base and is defined as follows:

'The 2001 Census has been conducted on a resident basis. This means the statistics relate to where people usually live, as opposed to where they are on Census night. Students and school children studying away from home are counted as resident at their term-time address. Wholly absent households were legally required to complete a Census form on their return. No information is provided on people present but not usually resident.'

The definition of the population used in the SOA estimates is consistent with the published local authority mid-year estimates, which is itself broadly consistent with the definition used for the 2001 Census.

In practice, when producing a population estimate, a number of data sources have to be used, each with its own definition of usual residence.

4. Mid-2001 SOA estimates

Findings of ONS research which reviewed evidence on the 2001 Census estimates indicated that, whilst no single piece of evidence on its own was conclusive, the weight of evidence suggested that the 2001 Census did not cover all people in England and Wales, particularly young adult men. Accordingly the 2001 local authority mid-year estimates, most recently revised in September 2004 reflect this evidence; these include adjustments for missing

Census forms, Longitudinal Study adjustments, the Manchester and Westminster Matching Studies and 2004 Local Authority Studies. The national count for mid-2001 is 318,000 greater than the 2001 Census count. Mid-2001 SOA population estimates have been produced to reflect the population adjustments incorporated in the local authority mid-year estimates, and will therefore differ to the published Census SOA estimates. The mid-2001 SOA estimates have been used as the population base for producing the mid-2002 SOA estimates, and in turn the mid-2002 SOA estimates become the base for the mid-2003 estimates etc.

Estimates of the resident population as at mid-2001 have been produced for publication by broad age group and sex for LSOAs and quinary age group and sex for MSOAs. For consistency, the mid-2001 LSOA population estimates, by age and sex, are constrained to the mid-2001 MSA estimates which in turn are constrained to the 2001 local authority mid-year estimates.

4.1 Methodology

The derived mid-2001 LSOA population base produced reflects the methodology used to create the revised mid-2001 national and local authority estimates, but down to the smaller LSOA geography. In summary, to produce the mid-2001 LSOA the stages below were followed:

1. Unadjusted Census Output Area counts by single year of age and sex for the usually resident population were aggregated to LSOAs.
2. To these counts, an adjustment was made for missing Census forms, adding 5,100 people in 21 LSOAs.
3. The population counts by single year of age and sex were aged forward from 29 April 2001 to 30 June 2001.
4. To these counts, the Longitudinal Study local authority population revisions to males aged 21-50 in the revised mid-2001 local authority revised estimates were added. Nationally 163,800 persons were added. These were disaggregated to 6,647 LSOAs, using an adaptation of the methodology used to derive local authority male counts. The methodology used to make the local authority revisions is available via the following link: [Click Here](#).
5. To these counts, 2,453 LSOA adjustments for 15 local authorities resulting from the local authority Studies review were added. Nationally this was 107,400 persons.
6. Births (0 year olds) occurring from 1 May 2001 to 30 June 2001 were added. Nationally this added 100,900 babies.
7. Deaths occurring from 1 May 2001 to 30 June 2001 were subtracted. Nationally this reduced the estimates by 84,100 persons.
8. The single year of age and sex counts were constrained to the revised 2001 local authority mid-year estimates (published in September 2004) and to reflect a population change of around 1,100 resulting from April 2003 boundary changes in three Welsh Unitary Authorities - Carmarthenshire, Ceredigion and Pembrokeshire.

MSOA estimates were produced by aggregating the LSOA counts. To address comments received following the user consultation with the previously published ward-level population estimates, further adjustments were incorporated into these LSOA/MSOA estimates as follows:

9. Adjustments for identified under/over estimation at MSOA level were applied to the component LSOAs. These adjustments reflect MSOAs where local authority Census response was poor and where the MSOA estimates following Longitudinal Study adjustments differed significantly to counts from administrative data sources. The adjustments resulted in some LSOAs having a reduction in population and others an increase. Net change at local authority level was zero. Adjustments were made to 818 LSOAs within 173 MSOAs in 13 local authorities. A list of these 13 local authorities is shown in Appendix B.
10. Due to inconsistencies between Census Home Armed Forces counts and mid-2001 Defence Analytical Service Agency (DASA) local authority counts, negative counts were produced by single year of age when the mid-2001 special population was deducted from the resident population for the production of mid-2002 estimates. To overcome this, adjustments were made for 51 LSOAs within 43 MSOAs in 33 local authorities which had negative counts when the special population was subtracted. Compensating adjustments were spread across all other LSOAs/MSOAs within these local authorities. Net change at local authority level was zero.
11. Due to inconsistencies between Census Prisoner counts and mid-2001 Home Office Prisoner counts, negative counts were produced by single year of age when the mid-2001 special population was deducted from the resident population for the production of mid-2002 estimates. To overcome this, adjustments were made for 6 LSOAs within 6 MSOAs in 6 local authorities which had negative counts when the special population was subtracted. Compensating adjustments were spread across all other LSOAs within these local authorities. Net change at local authority level was zero.

Further changes to the LSOA mid-2001 population were made to reflect changes in the allocation of postcodes to LSOAs. Postcodes which existed at the time of the 2001 Census which were split by electoral ward boundaries were assigned to a single Census Output Area on the basis of where the majority of the population lived. Other postcodes in use prior to April 2001, and new postcodes created after then, were assigned to a Census Output Area using the grid reference of the address closest to the mean of the easting and northing for each postcode. As a result some postcodes were allocated to a different Census Output Area than would have been the case if using a grid reference allocation. A new methodology was introduced with the August 2006 National Statistics Postcode Directory (NSPD) to resolve this.

All postcodes in England and Wales are now assigned by a point-in-polygon process (plotting grid references and assigning to digital boundaries by GIS) using current grid references. As the grid reference for each postcode will be current and the geography allocations will be directly based on it, the two will always correspond.

The implication of this change for small area population estimates is that the population residing within addresses for some postcodes will have a changed LSOA allocation, so postcodes included in a particular LSOA at the time of the 2001 Census will now be allocated to a different LSOA, thus affecting the mid-2001 base population estimates.

The postcodes and LSOAs which were affected the most in terms of overall population numbers were identified. Some LSOAs will have been over estimated at mid-2001 and others similarly under estimated. The postcodes investigated which had a changed LSOA allocation with the May 2006 NSPD were those which from the 2001 Census had a

population count of 100 or more, and which were previously assigned to a single Census Output Area, and where therefore the impact of possible adjustment would result in greatest improvement to the estimates. A 2001 Census population count of 100 was therefore the threshold considered for base population adjustment. 76 such postcodes were identified.

A visual check was then done using Ordnance Survey digital mapping to check that the new LSOA allocation was correct. In some cases it was identified that the original postcode to LSOA allocations were in fact more accurate. LSOA population adjustments for mid-2001 were made for 49 postcodes with an associated population count of 100 or more, this affected 88 LSOAs, within 62 MSOAs within 40 local authorities. For all of these postcodes the local authority allocation was unchanged.

Finally, in keeping with standard practice, the mid-2001 LSOA broad age estimates and MSOA quinary age estimates have been subject to disclosure control.

When constraining to the local authority mid-year estimates, special account was taken for the population change of around 1,100 persons (as at mid-2001) resulting from April 2003 boundary changes in three Welsh Unitary Authorities - Carmarthenshire, Ceredigion and Pembrokeshire. The SOA counts for these local authorities if aggregated to local authority level will not be consistent with the published local authority mid-year estimates, as these reflect the local authority boundaries as at mid-2001, whereas the SOA boundaries reflect the geography at the time of the April 2003 boundary changes.

Further local authority boundary changes took place in April 2005 affecting two Welsh Unitary Authorities - Neath Port Talbot and Powys with a population change of around 110 persons (as at mid-2005). These boundary changes were not reflected in the mid-2005 LSOA and MSOA estimates but have been reflected in subsequent estimates. As a consequence the mid-2006 to mid-2008 LSOA and MSOA estimates if summed will not be consistent with the published estimates for these local authorities.

4.2 Limitations with the estimates

It should be noted that no explicit adjustment has been made for either internal or international in- or out-migration in the nine-week period from Census day (29 April 2001) to 30 June 2001 (mid-year). However, an adjustment will have been made through the constraining to the local authority estimates which will have included these components. This therefore assumes that population change due to migration (in- and out-migrants) in this 9 week period in all SOAs within a local authority is proportional to its population size. This is unlikely to be true. However, as this estimate is for only 9 weeks it is unlikely that the differences will be large.

In addition, the distribution of the population adjustments made at the local authority level down to SOAs may not always accurately reflect the actual distribution of where the differences should be placed either geographically, and/or the age and sex distributions of the adjustments made.

The methodology used to create these mid-2001 SOA estimates has been quality assured and approved by ONS Methodology Directorate.

5. Mid-2002 SOA Estimates

The estimates for mid-2002 were produced using a Ratio Change methodology and have been revised to be consistent with the local authority mid-year estimates revised in August 2007. To produce these estimates, a number of different administrative datasets were used and these are documented in Appendix C.

5.1 Methodology

This describes the creation of the mid-2002 LSOA broad age group by sex estimates. The mid-2002 MSOA quinary age group by sex estimates were created in a similar manner using derived MSOA quinary age by sex change ratios.

The estimates were produced by applying the Ratio Change method to a LSOA estimate of the population base (the mid-2001 LSOA estimates) using a combination of patient registers, Child Benefit and Older Persons Datasets. Before applying these change ratios some population counts are subtracted (referred to as the special population) comprising UK armed forces, foreign armed forces and dependants, and prisoners, and added again after these counts are constrained to the 2002 local authority mid-year estimates minus the special population.

The main assumption behind this Ratio Change method is that, for each area, the data should have a consistent relationship with the true population over time.

Change ratios were calculated by quinary age group and sex for the Child Benefit data, patient register data and Older Persons Datasets. The change ratios are calculated by dividing for each dataset the mid-2002 count by quinary age and sex with the mid-2001 count by quinary age and sex, eg a mid-2002 count of 53 divided by a mid-2001 count of 48 gives a change ratio of 1.1042.

Some age groups are covered by a single dataset, while other age groups are covered by two of the three datasets. The datasets used to calculate the LSOA change ratios from mid-2001 to mid-2002 for quinary age groups and sex are as follows:

- 0-4, 5-9 and 10-14 - Child Benefit and patient registers
- 15-19 to 60-64 - patient registers
- 65-69 to 85+ - Older Persons Datasets and patient registers

In summary, these ratios are then:

1. applied to the mid-2001 LSOA population minus the mid-2001 special population by quinary age and sex, and
2. constrained to the mid-2002 MSOA estimates (less mid-2002 special population), which have been constrained to the local authority mid-2002 estimates less mid-2002 special population. Then,
3. mid-2002 LSOA estimates by single year of age and sex are produced by apportioning the quinary age counts to single year of age using mid-2002 local authority constrained patient register single year of age and sex counts;
4. these mid-2002 LSOA estimates by single year of age and sex are then constrained for consistency to mid-2002 MSOA estimates by single year of age and sex (these counts are derived from mid-2002 MSOA quinary age and sex estimates created using the same

ratio change methodology as for LSOAs, apportioned to single year of age and sex using mid-2002 local authority constrained patient register counts by single year of age and sex);

5. updated mid-2002 special population counts are then added back in to the quinary and single year of age and sex counts, to give mid-2002 LSOA estimates by quinary and single year of age and sex;

Estimates for mid-2002 LSOA broad age groups by sex are then created by aggregating the component quinary age group by sex estimates, and for the 0-15 and 16-29 year broad age groups, which cannot be derived solely from quinary age groups, through the additional use of single year of age and sex estimates.

In order to minimise the cell suppression for small numbers received from DWP for the Older Persons Datasets as part of disclosure control measures, data were received for an 80+ age group, rather than receiving separate counts of 80-84 and 85+. As our estimates have an upper age limit of 85+, the 80+ Older Persons Dataset counts were apportioned to 80-84 and 85+ according to the patient register counts.

Where two change ratios were produced for some of the age groups to reflect the availability of two datasets for these age groups, the change ratios were averaged by adding together the two ratios and dividing by two.

Any change ratios produced for counts containing a zero by quinary age and sex were changed to 1 to ensure a valid ratio was produced. Where change ratios were applied to a base population by quinary age and sex of zero, the base population was changed to 1 to ensure an actual population count in the estimate, otherwise counts of zero in the base population would forever remain at zero.

Where SOA data for any of the datasets were identified to be erroneous or extremely unrealistic, the calculated change ratios were changed (typically to 1), or changed to the change ratio for the other dataset for a particular age group. Criteria were developed to assist in the consideration of making any changes to the originally calculated change ratios.

For the LSOAs and MSOAs, the following ratio changes were made:

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	323	0.94	59	0.79
Older Persons Dataset	108	0.31	27	0.38
Patient Registers	13	0.04	11	0.15

The changes to the Child Benefit ratios reflect large clusters of LSOAs/MSOAs within certain local authorities where there were known deficiencies in the data for mid-2001, namely Tewkesbury in Gloucestershire; Three Rivers and Watford in Hertfordshire; Blyth Valley in Northumberland; Cherwell and West Oxfordshire in Oxfordshire; and Monmouthshire.

As with the mid-2001 SOA estimates, the SOA counts in Carmarthenshire, Ceredigion and Pembrokeshire will not be consistent with the published local authority mid-year estimates, as the local authority estimates reflect the local authority boundaries as at mid-2002, whereas the SOA estimates reflect the local authority boundaries as at May 2003.

In keeping with standard practice, the mid-2002 LSOA and MSOA estimates have been subject to disclosure control. An illustrative diagram of the Ratio Change method is shown in Appendix D.

5.2 Limitations with the estimates

The estimates have been produced using administrative data sources to identify annual population change. Any deficiencies in these data sources may therefore impact upon the quality of the estimates produced. Where known deficiencies have been identified, corrective measures have been applied, though other deficiencies in the use of administrative data sources for producing population estimates may be less apparent, for example list cleaning of the patient registers. In addition there is no intention to imply accuracy with the LSOA estimates by broad age group and sex, and the MSOA estimates by quinary age group and sex down to unit level, but simply that it is more appropriate to publish the estimates in this way to ensure consistency between counts by age group and total counts.

6. SOA estimates mid-2003 onwards

The same Ratio Change methodology has been used to produce SOA estimates for mid-2003 onwards as was used for the mid-2002 estimates. The mid-2002 SOA estimates were used as the base population from which change ratios for mid-2002 to mid-2003 were applied to create the mid-2003 estimates, whilst the mid-2003 SOA estimates were used as the base population from which change ratios for mid-2003 to mid-2004 were applied to create the mid-2004 estimates etc.

The estimates for mid-2003 onwards are consistent with the local authority mid-year estimates.

On an annual basis, the following ratio change adjustments were made:

Mid-2003

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	17	0.05	0	0
Older Persons Dataset	33	0.10	12	0.16
Patient Registers	6	0.02	1	0.01

Mid-2004

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	13	0.04	1	0.01
Older Persons Dataset	47	0.14	10	0.14
Patient Registers	19	0.06	16	0.22

Mid-2005

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	14	0.04	0	0
Older Persons Dataset	25	0.07	6	0.08
Patient Registers	13	0.04	12	0.16

Mid-2006

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	20	0.06	2	0.03
Older Persons Dataset	N/A		N/A	
Patient Registers	15	0.04	15	0.21

Mid-2007

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	43	0.13	0	0
Older Persons Dataset	N/A		N/A	
Patient Registers	12	0.03	16	0.22

Mid-2008

Dataset	LSOAs		MSOAs	
	No.	%	No.	%
Child Benefit	N/A		N/A	
Older Persons Dataset	N/A		N/A	
Patient Registers	12	0.03	25	0.35

These ratio change adjustments all reflect occurrences where the implied ratio from the actual administrative dataset counts, if used alone or averaged with a change ratio from another dataset (depending on the age group), would give rise to implausible population estimates when applied to the previous years' population estimate.

Please note that at the time of the production of the mid-2008 estimates, the Child Benefit data had not been received and therefore the change ratios for the population aged 0 to 14 were derived solely from the patient register data. In addition, at the time of the production of the mid-2006, mid-2007 and mid-2008 estimates, the Older Persons Dataset had not been received and therefore the change ratios for the population aged 65 and over were derived solely from the patient register data.

7. Future changes to methodology

The SOA estimates are initially released on an experimental basis. While estimates are being published on this basis, all aspects of the methodology together with other identified shortlisted methods are subject to further investigation to see if quality improvements can be made. User feedback on these estimates is very much welcome, as this may assist in improving future estimates.

8. Further Information and Contacts

Information on the Small Area Population Estimates project for England and Wales is available on the ONS website: <http://www.statistics.gov.uk/sape>

For user feedback and further information, please contact the Subnational Statistics Unit:

Email: SAPE@ons.gov.uk

Telephone: 01329 444664

Appendix A

Shortlisted Methods

Following an evaluation of identified methods against a set of evaluation criteria, the following three methods were shortlisted - Apportionment, Cohort Component and Ratio Change. These were all evaluated in order to identify a preferred method, as part of the evaluation of each of the methods; consideration was also given to the capability of each method to make use of new data sources should they become available. The common datasets for all these methods are described in Appendix C.

Apportionment method

This method breaks down the 'known' population of a larger area (eg local authority mid-year estimates) into smaller areas using an indicator of population for the small areas, for example patient register data. Inherent in this method is the assumption that the relationship between the indicator of population and the true population is the same for each small area within the reference area. Apportionment requires data only for the current time, no base population is required and is therefore straightforward to use even in areas where there have been recent boundary changes. If the data source used for apportioning has an age and sex breakdown then estimates by age and sex may be produced. The data sources used for this method are: patient registers, Child Benefit, Older Persons Datasets, births, infant deaths, UK and foreign armed forces, foreign armed forces dependants, prisoners and school boarders.

Cohort Component method

Cohort Component or cohort survival can be referred to any method that involves the ageing of the population age structure which has been estimated for an earlier year, and the use of birth counts to estimate the number of infants and an allowance for deaths and migration. The assumptions necessary using this method depend on how the components of population change: births, deaths and migration, are estimated. This method is used to produce the local authority mid-year estimates. The data sources used for this method are: patient registers (for internal migration), births, deaths, UK and foreign armed forces, foreign armed forces dependants and prisoners. Unlike the local authority mid-year estimates, when using the Cohort Component method to produce small area population estimates, no explicit estimation of international migration has been made. A method to estimate this international migration at small area level was tested, but did not prove to be successful.

Ratio Change method

An indicator of the current and the past population is used to update an earlier population estimate for each small area. The assumption is that the relationship between the indicator of population and the true population has remained the same for the small area since the base year or latest estimated year. For example the population aged 0 to 4 in an SOA may be estimated as the population aged 0 to 4 in the base year, times the Child Benefit counts of 0 to 4 year olds for the current year divided by the Child Benefit count of 0 to 4 year olds in the base year. The data sources used for this method are: patient registers, Child Benefit, Older Persons Datasets, UK and foreign armed forces, foreign armed forces dependants and prisoners.

Appendix B

Local Authorities receiving Population under/over estimation adjustments with mid-2001 estimates

Local Authority	No. LSOAs	% all LSOAs in LA	Largest adjustment (+ or -)	No. MSOAs	% all MSOAs in LA	Largest adjustment (+ or -)
Brent	94	54.0	+456	18	52.9	+2,486
Camden	51	38.0	-176	11	39.3	-823
Hackney	68	49.6	-417	14	51.9	-1,597
Hammersmith & Fulham	42	37.8	+150	10	40.0	+563
Haringey	64	44.4	+392	16	44.4	+1,145
Lambeth	118	66.7	-326	23	65.7	-1,816
Lewisham	64	38.6	+235	14	38.9	+978
Newham	52	32.7	+144	12	32.4	+941
Tower Hamlets	48	36.9	-264	11	35.5	-934
Manchester	97	37.5	-558	20	37.7	-748
Nottingham	40	22.7	-217	8	21.6	-1,212
Slough	43	55.1	+266	8	57.1	+1,263
Southampton	37	25.3	+401	8	25.0	+1,336
Total	818			173		

In the previous estimates prior to the October 2007 revisions, adjustments were also made in Kensington & Chelsea, Southwark and Westminster; these adjustments have been reversed as a result of further investigation.

For those LSOAs or MSOAs receiving positive adjustments by quinary age and sex (ie more people added), the sum of these adjustments within a local authority will equal the absolute sum of the negative adjustments in the other LSOAs or MSOAs (ie where people have been deducted). For the local authority as a whole there will therefore be a net quinary age and sex adjustment of zero.

Appendix C

Administrative datasets used

Following an evaluation of a number of data sources, the following administrative datasets were used with the Ratio Change method to produce the mid-2002 to mid-2005 SOA estimates:

Patient Registers

Year specific July counts of individuals included on the NHS patient registers were used, ie persons registered with a doctor, by single year of age and sex at postcode level. Through the use of a postcode look-up table (eg the National Statistics Postcode Directory) these counts can be aggregated to different geographical levels such as wards and SOAs.

The data provided to ONS have already been validated, so only records with a valid postcode are received. Improvements in the quality of postcodes on the patient registers reflect the efforts by the Strategic Health Authorities (in England) and Health Boards (in Wales) to improve the quality of the data on their registers.

Previous ONS research², describing the use of patient registers to measure internal migration in England and Wales concluded that, at the local authority level, patient registers counts do not provide a reliable estimate of the resident population of England and Wales. It was also noted that for the patient registers to be used in this way the counts would require significant adjustments and further research would be needed.

When the patient register count exceeds the mid-year population estimates, this is often referred to as list inflation. This may occur when some patients have more than one NHS number and are double counted, and patients may be on doctors' lists after having left the country. List inflation may also be localised, for example in student areas where students do not quickly re-register after finishing their course of study and moving away from an area.

List inflation is greatest in London boroughs. It is expected that this list inflation is due mainly to a high number of international migrants moving into London and registering with a GP³. Many of them will not be removed from the GP lists when they leave.

Conversely there are other areas where the patient registers will be missing individuals due to these persons being ineligible to be registered with a GP, this has predominantly been in areas where there is known to be a high number of armed forces. This group has not generally been registered with a NHS GP. Other groups of the population which are excluded from the NHS are prisoners that are sentenced for a term of two years or more and certain patients in long stay medical hospitals. In addition there are individuals who obtain all their medical care privately who may not be registered with a GP, but these numbers are likely to be small.

An investigation has been undertaken for individual postcodes where there are large patient register counts and instances where the patient register counts fluctuate significantly over

² Scott A and Kilbey T (1999). Can Patient Registers give an improved measure of internal migration in England and Wales? *Population Trends* 96, pp 44-55.

³ Office for National Statistics (2002). Using Patient Registers to estimate Internal Migration – Customer Guidance Notes. Migrations Statistics Unit, ONS, August 2002.

time. It emerged that a significant number of these postcodes related to halls of residences at Universities. Through data from Universities and the 2001 Census, for some areas an adjusted patient register postcode count was used to more closely reflect actual numbers of students or other residents. Unless there was information to suggest otherwise, these adjusted patient register postcode counts by age and sex are kept constant over time. Around 470 postcodes are adjusted in this way, reducing the overall patient register count by around 80,000.

Child Benefit data

The Department for Work and Pensions (DWP) previously administered a Child Benefit database containing details about children for whom Child Benefit was claimed, along with the claimant's particulars (usually the mother). The claimant's details were therefore repeated in as many records as she/he has children. In 2003 responsibility for Child Benefit passed to Inland Revenue, now HM Revenue and Customs.

Child Benefit counts for April 2001 and August 2002 were received from DWP, June 2003 counts from Inland Revenue, and August counts for subsequent years from HM Revenue and Customs (except for 2008, where data have not been received). Mid-year counts were not available. Whilst data were held by DWP/IR/HMRC for each child for whom Child Benefit was claimed, because of social security and data protection legislation, DWP/IR/HMRC were unable to give ONS access to identifiable, subject level data. Through an arrangement with DWP and ONS, the University of Oxford carried out a LSOA level aggregation of the 2001 and 2002 datasets. Whilst Child Benefit may be claimed for children aged 0 to 16 and over, only counts for children in the quinary age groups to 0 to 4, 5 to 9 and 10 to 14 were used to calculate change ratios. Eligibility for benefit decreases for children older than 15.

There are a number of valid reasons why we would expect the count of children aged 0 to 14 from Child Benefit to be lower than the national mid-year estimates, these include:

- Dependants of students and foreign nationals, including foreign armed forces are not eligible for Child Benefit
- Children in local authority care or foster care are not eligible for Child Benefit
- Children detained in secure or non-secure accommodation are not eligible for Child Benefit
- Children whose entry to the United Kingdom is subject to immigration control are not eligible for Child Benefit
- Children for whom Child Benefit is not claimed but who are eligible

There are also valid reasons why we would expect the distribution (location) of children from Child Benefit to differ to the mid-year estimates, these include:

- School boarders – where claimant's address is different to boarder's residential address
- Children who reside at a different address to the address of the claimant

Older Persons Dataset

DWP Older Persons Datasets were used for August 2001 to August 2005, but have not been provided for 2006, 2007 or 2008. The 2001, 2002, 2004 and 2005 datasets were created by the University of Oxford from individual DWP databases for Attendance Allowance, Disability Living Allowance, Widows Benefit, State Pension, Incapacity Benefit, Winter Fuel Payment and Minimum Income Guarantee, and covers persons aged 65 years and over. DWP processed and supplied the 2003 dataset using the same DWP benefit databases used by

the University of Oxford for the other years. It is estimated that the number of people who are not eligible for any of the above benefits is relatively small and that of those who are eligible for State Pension, the take-up is thought to be very high.

Because of social security and data protection legislation, DWP is unable to give ONS access to identifiable, subject level data and so a similar arrangement as with the Child Benefit data was set up with the University of Oxford to provide us with aggregated LSOA counts of persons aged 65 years and above for 2001, 2002, 2004 and 2005.

UK Armed Forces (special population)

The armed forces component of the local authority usually resident population estimates includes UK armed forces, covering personnel in the air force, army and navy stationed in England and Wales. The numbers are collected annually from the Defence Analytical Services Agency (DASA) who provides data on the number of UK armed forces by gender stationed at each base in England and Wales. These data are mainly based on administrative systems that are used for pay and personnel purposes.

DASA also provides an age and sex breakdown for all UK armed forces which is used to help derive an age and sex breakdown of armed forces personnel at local authority level. DASA are currently unable to provide counts of armed forces personnel on a usually resident basis, residence information is available from the Census.

For England and Wales, the number of UK armed forces from the 2001 Census was 154,000 compared with the DASA mid-2001 count of 150,900, a difference of 3,100 (2.1 per cent). Some of these differences may be due to the timing differences between the Census and mid-year. It is not generally feasible to disaggregate the UK armed forces data currently available at local authority level and used within the local authority mid-year estimates to SOA level without using a Census SOA distribution. However where information is known about changes to the location of Armed Forces personnel, this is reflected in the estimates.

Foreign Armed Forces and dependants (special population)

The foreign armed forces component of the local authority usually resident population estimates covers personnel in the US air force, US army and US navy, together with their dependants. Personnel from the US air force make up the great majority of foreign armed forces stationed in England and Wales.

The age and sex distribution of these personnel is not available annually and an age and sex distribution is applied using Census data. The data sources for the three armed forces services differ and the format of data received differs slightly between sources. Data on US air forces personnel and dependants for example are only available by location (town) whilst data on US army personnel and dependants are available at postcode level and data on US navy personnel and dependants are available at postcode sector level.

For England and Wales, the number of foreign armed forces and dependants from the supplied mid-2001 data was 19,300 compared with the 2001 Census count of 17,600, a difference of 1,700 (9.4 per cent). This difference may be due in part to the timing differences between the Census and mid-year. It is not currently feasible to disaggregate the foreign armed forces data currently available at local authority level and used within the local authority mid-year estimates to SOA level without using a Census SOA distribution.

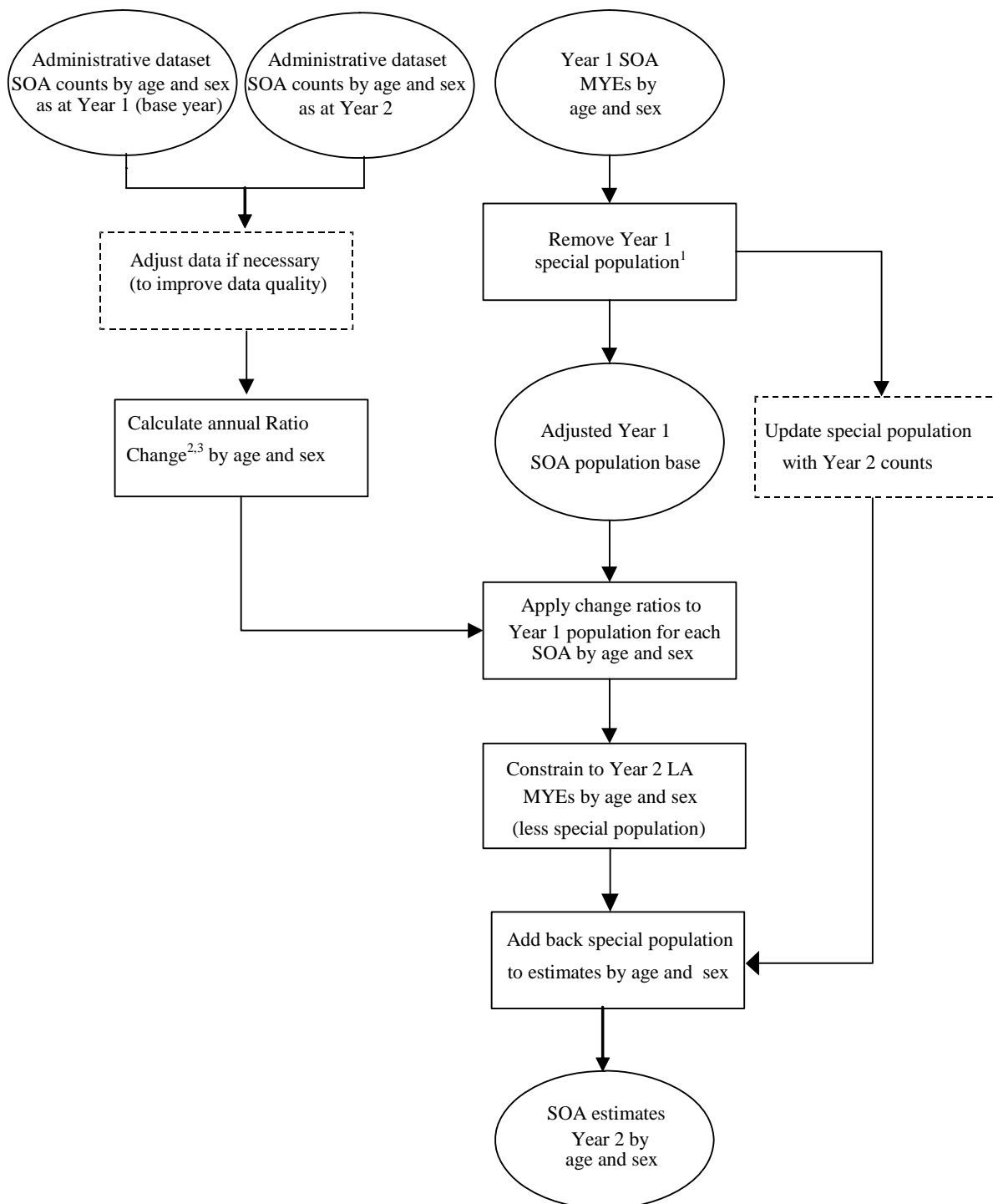
Prisoners (special population)

Data are received from the Ministry of Justice (formerly the Home Office) on the number of prisoners for inclusion in the local authority mid-year estimates. For the purpose of the local authority mid-year estimates, prisoners are regarded as usually resident in a prison if they have been sentenced and have served 6 months or more of their sentence. Age and sex information for each prisoner for all prisons is received.

The count of prisoners included in the 2001 local authority mid-year estimates was 36,100 which compares to a count of 47,000 in the 2001 Census, a difference of 10,900 (30.2 per cent). However there are definitional differences between these counts. The local authority mid-year estimates are a subset of the Census count which included those persons held in prison on Census night who had no other usual residence or who would not be included on a Census form elsewhere.

Appendix D

Simplified Ratio Change methodology diagram of production of SOA Estimates



1. Special Population adjustments are as follows (affecting some age groups only):
UK armed forces, foreign armed forces and dependants, and prisoners
2. Ratios produced for counts containing a zero by age and sex are changed to 1

Example of a Ratio Change calculation by quinary age and sex:

Year 2 count of 226 ÷ Year 1 count of 197 gives a change ratio of 1.1472 which is applied to the Year 1 population for the appropriate age and sex group

For some age groups averaged ratios by sex are used eg Males 0-4:
(patient register Male 0-4 ratio + Child Benefit Male 0-4 ratio) ÷ 2