

# **Evaluation of the Census 2011 Quality Assurance Studies project**

**December 2010**

## **Contents**

1. The quality assurance (QA) studies project
2. Overview of census data quality assurance process
3. QA studies
4. Council tax versus address register analysis
5. Average household size analysis
6. Electoral registration analysis
7. Houses in Multiple Occupation
8. Housing benefit
9. Quality templates
10. Supplementary evidence
11. Lessons learnt
12. Recommendations
13. Continuing engagement with local authorities

# 1. The QA studies project

## 1.1 QA objectives

Engagement with local authorities (LAs) about census quality assurance has two key objectives:

- to improve LAs' understanding of, and build confidence in, the census results by informing them about the quality assurance process and the challenges faced
- to develop the best possible understanding of each LA's population ahead of the census, with clear evidence about where there are discrepancies/concerns in the mid-year estimates. This means identifying and securing access to local LA sources where appropriate

As was the case in 2001, there are two issues that will drive LAs to query the 2011 Census estimates:

- any significant unexplained divergence in the 2011 Census estimates, particularly lower population estimates than the latest mid-year estimates (e.g. 2010); and
- where LAs believe that their current mid year estimates (MYEs) and/or the 2001 Census estimates are too low based on evidence from other data sources

The first issue will be identified during the quality assurance of the census estimates. The checks conducted will seek to explain these differences using administrative sources, demographic analysis and intelligence from the census operation. Differences will inevitably occur. There are, for example, definitional and coverage reasons why administrative sources will give different estimates to that measured by the census.

The second issue can be identified ahead of the census by engaging with LAs early to understand their concerns and the data sources that LAs draw on to question particular components of their mid-year estimates and/or census estimates. Preparation for the 2011 Census includes compilation of correspondence between ONS and LA users since 2001 (including correspondence on the 2001 Census estimates). Understanding LAs' concerns and evidence will help ONS to improve the census quality assurance process, ultimately resulting in the publication of census population estimates in which LAs have confidence.

## 1.2 Background to QA studies

Experience suggests that the census produces population estimates that are fit for purpose in the vast majority of LAs. However in some LAs the demographic composition of the area, or societal changes since 2001, suggest that additional research to understand the complex issues in these areas is required. It was therefore proposed that ONS undertakes a series of QA studies working closely with some of these LAs to further understand their concerns and how the QA process can be improved for all LAs. It was also intended to use the QA studies to identify data sources held by LAs which will be of benefit to the QA process and where appropriate to gain access to these datasets.

On 27 October 2009, the Census Regional Champions Meeting endorsed a proposal for ONS to undertake a pilot for the QA studies project to trial the approach. This pilot involved ten LAs and ran between February and April 2010. Following the successful completion of these studies a wider QA study was undertaken between July and October 2010. A further 30 LAs were involved in this work.

This report gives an overview of the engagement with local authorities, the analysis undertaken and the key findings and recommendations from the QA studies project.

## 2. Overview of census data quality assurance process

The 2011 Census has four strategic aims:

- to provide high quality statistics that meet user needs
- to build confidence in the final results
- to provide value for money solutions, and
- to protect, and be seen to protect, confidential personal census information

Quality throughout the census operation will be managed using a quality model (defined in the ONS 2011 Census Quality Strategy<sup>1</sup>) that involves:

- design quality
- operational quality management
- quality assurance
- quality measurement and reporting
- the production of high quality population statistics

The Data Quality Assurance Strategy<sup>2</sup> contributes to the quality model by validating census data from the period before census day, when management information and early census returns will provide early evidence of response patterns and characteristics, through to the publication of outputs.

The key objectives of the data quality assurance work outlined in the Data Quality Assurance Strategy are:

- to ensure that 2011 Census outputs are fit for purpose and meet user expectations
- to be able to understand differences between census population estimates and rolled-forward mid-year population estimates or other survey and administrative sources and explain these i) to the QA Panel, which has responsibility for recommending approval or rejection of each LA's census population estimates, and ii) to census users through the production of informative metadata
- to ensure that census information on population structures and characteristics is accurate
- to be transparent in the methods and implementation of data QA
- to plan and implement QA activities in partnership with census stakeholders
- to provide indicators of census quality in Quality Reports that will be released at the same time as census outputs
- to provide timely census QA input to the production of 2011 mid-year population estimates in 2012

The accuracy of census returns will be assessed by comparing census distributions against equivalent measures from ONS products and administrative and survey sources using a series of QA checks. The main focus will be the quality assurance at LA level. These checks will be undertaken at a number of stages throughout census processing to identify potential systematic error or bias in returns. A list of admin sources that will be used in the QA process or which are still being investigated are shown in Annex A.

For the 2011 Census QA, more extensive use will be made of demographic analysis including sex ratios, fertility and mortality rates which will be used in combination with other evidence to indicate where there may be problems with the age by sex counts.

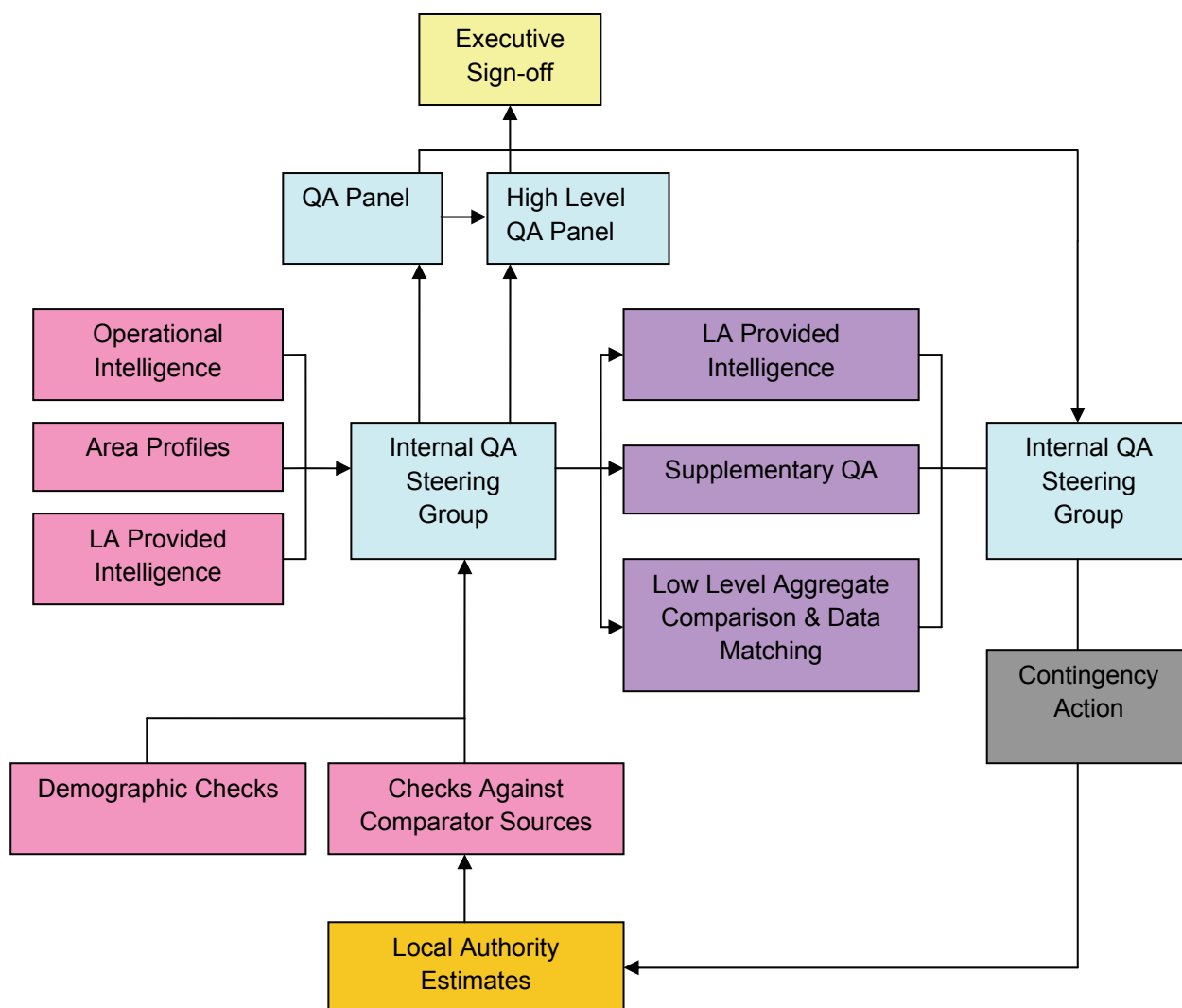
---

<sup>1</sup> The ONS 2011 Census Quality Strategy is available at: <http://www.gro-scotland.gov.uk/files1/stats/census-quality-strategy-2011.pdf>

<sup>2</sup> The 2011 Census Data Quality Assurance Strategy is available at <http://www.ons.gov.uk/census/2011-census/process-info/data-quality-assurance/index.html>

Figure 1 gives an overview of the process that will be used to quality assure the local authority census estimates.

**Figure 1**



### 2.1 Automated quality assurance

All local authority level estimates will be subject to a series of automated checks. These checks will be pre-programmed and will initially be assessed at the local authority (LA) level. The automated checks will form the basis of a QA pack for each LA, which will be presented to the QA panel to inform their decisions on the acceptance or rejection of census estimates.

The high priority automated checks are listed below:

QA check	Comparator dataset	Check approach
1. Age and sex	<ul style="list-style-type: none"> <li>• Birth registrations</li> <li>• Patient register</li> <li>• Mid-year population estimates</li> </ul>	<ul style="list-style-type: none"> <li>• Quinary age and sex</li> <li>• Absolute values and proportions</li> <li>• Diagnostic range<sup>3</sup> capped to avoid implausibly large ranges</li> </ul>

<sup>3</sup> The diagnostic range method is used for checks with more than one comparator data source to set upper and lower bounds within which census estimates are expected to fall. The method subtracts the lowest value

	<ul style="list-style-type: none"> <li>• School Census</li> <li>• Central Information System</li> </ul>	
2. Household Number and Average Size	<ul style="list-style-type: none"> <li>• Council tax</li> <li>• Address register</li> <li>• Patient register</li> <li>• Communities and local government household estimates</li> <li>• Welsh Assembly Government household estimates</li> </ul>	<ul style="list-style-type: none"> <li>• Household number by occupied and non-occupied. Identification of dummy form<sup>4</sup> responses</li> <li>• Distribution of average household size</li> <li>• Tolerance methods currently under development</li> </ul>
3. Ethnicity	<ul style="list-style-type: none"> <li>• Population Estimates by Ethnic Group</li> <li>• Integrated Household Survey</li> <li>• School Census</li> <li>• Independent Schools data from Department for Education and Welsh Assembly Government</li> </ul>	<ul style="list-style-type: none"> <li>• Broad ethnic group by sex for all ages</li> <li>• Detailed ethnic group by sex for all ages</li> <li>• Broad ethnic group by sex and quinary age group</li> <li>• Detailed ethnic group by sex and quinary age group</li> <li>• Absolute values and proportions</li> <li>• Diagnostic range where multiple comparators allow</li> <li>• Tolerances based on School Census data only are currently underdevelopment</li> </ul>
4. Students	<ul style="list-style-type: none"> <li>• Higher Education Statistics Agency (HESA)</li> <li>• Further Education Student Numbers from Business, Innovation and Skills</li> </ul>	<ul style="list-style-type: none"> <li>• Students over 18 by age and sex</li> <li>• Students in communal establishments</li> <li>• Tolerances relative to the quality of HESA data and the number of further education students as a proportion of the LA population</li> </ul>
5. Armed Forces (Home/Foreign)	<ul style="list-style-type: none"> <li>• Defence Analytical Services Agency</li> <li>• US Armed Forces</li> </ul>	<ul style="list-style-type: none"> <li>• Age and sex based checks</li> <li>• Tolerance methods currently being developed</li> </ul>
6. Migration (Internal)	<ul style="list-style-type: none"> <li>• Patient registration</li> </ul>	<ul style="list-style-type: none"> <li>• Age and sex based checks</li> <li>• Tolerance methods currently being developed</li> </ul>
7. Migration (International)	<ul style="list-style-type: none"> <li>• Patient registration</li> <li>• International Passenger Survey</li> <li>• National Insurance numbers</li> </ul>	<ul style="list-style-type: none"> <li>• Age and sex based checks</li> <li>• Tolerance methods currently being developed</li> </ul>
8. Large Communal Establishments	<ul style="list-style-type: none"> <li>• Ministry of Justice prisoner numbers</li> <li>• Higher Education Statistics Agency (HESA)</li> <li>• Patient registration</li> <li>• Questionnaire Tracking System<sup>5</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Number of residents by age and sex</li> </ul>

comparator from the highest to establish a range which is then applied either side of the midpoint between the highest and lowest values. Census estimates falling outside of this range are flagged for further investigation.

<sup>4</sup> A dummy form is a placeholder questionnaire completed by field staff for households where they have been unable to collect a return. Dummy forms contain basic information about the household.

Quality Assurance IT systems will provide functionality that will facilitate the QA process. This will include:

- a processing summary to identify the current status of LAs and the stages for which checks are available
- the flagging of census estimates as red or green to allow users to quickly identify areas where census values are outside expected ranges and where additional analysis is required
- 'drilldown' functionality to explore the data below LA level (down to output area where data allows) and within variable categories, for example by detailed ethnic group, having first assessed broad ethnic groups
- cumulative analysis – Allows the Census Quality Team to look at census distributions against comparators for aggregate groups of LAs. These groups can be standard administrative geographies, for example Government Office Regions or bespoke groups of LAs. For example, it might be useful to combine the data for all LAs which contain university halls of residence
- cross checking – This allows users to compare census estimates across different processing stages, for example to assess the impact on the estimates before and after coverage estimation<sup>6</sup>

## 2.2 Supporting evidence

In addition to the checks shown above, the Census Quality Team will have access to a range of other information to help interpret census estimates and to identify where further analysis is required. The information will include:

- early extract data – A daily extract of all census returns coded that day starting two weeks prior to census day. This will allow early assessment of the census data prior to the main data deliveries and will help to identify systematic processing errors. It will also provide an early indication of respondent error and bias
- area profiles – A summary of the area's population, including data from 2001 and changes since the last census will be available to the QA team for consideration. This information has been pre-specified and highlights any potentially difficult to count populations
- operational intelligence - The QA process will consider information from census and Census Coverage Survey (CCS) field operations, such as return rates and evidence of localised field failure and information from the Questionnaire Tracking System and Census Management Information System. Diagnostics will also be available from the processing stages through which census data will pass
- LA provided evidence - A range of LA provided evidence will be available for consideration. This includes information/analysis provided from those LAs involved in the QA studies and communication with ONS on its mid-year population estimates since the last census. The information that has been provided by LAs is recorded in their Census Local Partnership Plans (CLPPs)

---

<sup>5</sup> The Questionnaire Tracking System contains information on all residential addresses in England and Wales. It records the current status of addresses and any questionnaires associated with them. During the operational period new addresses may be added as a result of either fieldwork or through requests to the Contact Centre. The QT system also includes estimated numbers of residents in communal establishments.

<sup>6</sup> Coverage estimation – A Census Coverage Survey (CCS) is carried out 6 weeks after census day to estimate the level of people and households not counted in the census. Statistical modelling techniques are used to produce local authority estimates which account for the level of undercount identified via the CCS whilst also allowing for estimated overcount. More information on this process can be found at <http://www.ons.gov.uk/census/2011-census/process-info/statistical-meth/coverage-assessment-leaflet.pdf>

The automated checks will be reviewed alongside any supporting evidence. This information will form the basis of a QA pack for each LA.

### 2.3 Supplementary QA

Where there are specific issues or areas of concern highlighted within the automated checks, the team will undertake supplementary QA. The aim of this analysis is to further explore, explain and, in some cases, adjust the census estimates. For example, if an LA is showing an unusual fertility pattern compared to historic trends, an analysis of fertility by country of birth might help explain why rates are different to expectations.

Supplementary QA is not routinely run for all local authorities. It will draw on a wide range of admin and survey data sources, including (in some cases) datasets for a specific area provided by the local authority as part of the LA engagement work prior to the census.

### 2.4 Low level aggregate comparison and data matching

A further extension to the supplementary QA will make use of low level aggregate comparisons of census data with administrative sources, for example at postcode level. Where administrative sources are available at record level, there is potential to use data matching. Such analysis will be carried out where unexplained inconsistencies remain from previous analysis.

### 2.5 Potential actions

There are a number of actions which might be taken if the quality assurance activities identify an issue. Options include:

- adjustment based on other information collected in the census, for example on visitors or second residences
- amending the coverage adjustment process based on evidence from alternative data
- calibration of estimates to external data either locally or nationally

A paper outlining the QA approach, including improvements, will be published ahead of the census.

## **3. QA studies**

In this report 'QA studies' refers to the work undertaken for both the pilot and wider group of LAs. Local authority names have been removed and detailed data excluded from the analyses presented to ensure information is not disclosive.

### 3.1 Aims and objectives of the QA studies

The QA studies project had several aims:

- identify LA data sources that could serve as comparators during QA analysis
- improve ONS's understanding of the quality of comparator data sources for QA
- provide softer intelligence on population sub-groups, unique population characteristics or any other issue that could help ONS's understanding of the census results
- establish relationships between ONS and LAs to support the QA work
- inform process / engagement required for LAs not involved in the pilot or QA studies

### 3.2 Selecting local authorities for the QA studies



The selection of LAs invited to participate in the QA studies was based on (i) those which have experienced high migration since 2001 and (ii) those where enumeration is anticipated to be most difficult based on estimated non-response modelling work undertaken by ONS Methodology.

Ten LAs were selected for the pilot. It was necessary to limit the number of London Boroughs involved to four to ensure the pilot had a broader geographic coverage. For the wider QA studies work, 41 LAs were invited to participate, of which 30 accepted (shown below). Geographic coverage was maintained by inviting a minimum of 3 LAs per region.

<b>EAST MIDLANDS</b>  <b>Pilot:</b> Nottingham, <b>QA studies:</b> Derby, Leicester, Lincoln	<b>SOUTH EAST</b>  <b>Pilot:</b> Oxford <b>QA studies:</b> Guildford, Slough
<b>EAST OF ENGLAND</b>  <b>Pilot:</b> Cambridge <b>QA studies:</b> Forest Heath, Luton	<b>SOUTH WEST</b>  <b>Pilot:</b> n/a <b>QA studies:</b> Bournemouth, Bristol, Exeter
<b>LONDON</b>  <b>Pilot:</b> Camden, Hackney, Kensington & Chelsea, Westminster <b>QA studies:</b> Hammersmith & Fulham, Harrow, Hounslow, Islington, Lambeth, Newham, Southwark, Tower Hamlets, Wandsworth	<b>WALES</b>  <b>Pilot:</b> Ceredigion <b>QA studies:</b> Cardiff, Newport, Swansea
<b>NORTH EAST</b>  <b>Pilot:</b> Newcastle upon Tyne, <b>QA studies:</b> Gateshead, Middlesbrough, South Tyneside	<b>WEST MIDLANDS</b>  <b>Pilot:</b> n/a <b>QA studies:</b> Birmingham, Warwick
<b>NORTH WEST</b>  <b>Pilot:</b> Manchester <b>QA studies:</b> Liverpool	<b>YORKSHIRE &amp; THE HUMBER</b>  <b>Pilot:</b> n/a <b>QA studies:</b> Leeds, Sheffield

### 3.3 QA studies approach

ONS followed a consistent approach for the pilot LAs and wider group of QA studies areas. For each exercise ONS:

- held an initial meeting to discuss the aims and objectives of the project and the sources of data available within LAs
- made a request for data and supplementary evidence
- analysed the data received
- shared the findings of the research by producing an analysis pack for each participating LA

For the pilot LAs, ONS presented the key findings of the analyses and distributed the LA packs at a meeting in April 2010. This allowed ONS to seek feedback on potential reasons for the trends and

anomalies identified in the analysis. It also allowed ONS to evaluate the pilot, review lessons learnt and agree the approach for the wider group.

### 3.4 QA studies analysis

A meeting was held with representatives from all pilot LAs on 25 January 2010. The main focus of the discussion was on sources of data available within local authorities which could add value to the QA process. LAs asked that ONS be very clear on exactly which data sources they should provide rather than leaving them with an open-ended request. They gave a clear steer on the potential for the following sources:

- Council tax data
- Electoral registration data
- Patient register data (microdata for England & Wales held by ONS)
- Housing benefit data
- Houses in Multiple Occupation (HMO) data held locally

It was noted that the census QA plans already propose the use of many of these datasets. It was agreed that the pilot LAs would provide these data to ONS, with the exception of the patient register data which it already holds. ONS committed to undertake some analytical work using the data received in conjunction with sources already held within ONS. The aim of the analysis was to gain a better understanding of the quality and coverage of each source and its suitability as a comparator for census quality assurance.

A range of other sources were also discussed including:

- School Census data (microdata for England held by ONS)
- Independent schools data
- Driver and Vehicle Licensing Agency (DVLA) data
- Companies House data
- Residents parking data
- Concessionary fares

A summary of the analysis carried out by ONS is given below. An evaluation of the pilot determined which analyses there was value in repeating or extending for the wider QA studies.

<b>Analysis</b>	<b>Pilot LAs</b>	<b>QA studies</b>
Council tax versus address register <sup>7</sup> analysis using council tax exemption and discount information to explain differences between the two sources	✓	✓
Average household size analysis using patient register data as an estimate of the number of people and council tax data as an estimate of the number of households by Lower Super Output Area (LSOA) <sup>8</sup>	✓	✓

<sup>7</sup> An address register is being compiled by the ONS to produce a postal list for 2011 Census questionnaires. This address register is produced by matching the existing address intelligence from the National Land & Property Gazetteer (NLPG), the Royal Mail's Postal Address File (PAF) and the Ordnance Survey's Address Layer 2 (AL2). A proportion of the unmatched addresses are sent to the address suppliers, local authorities and the ONS team undertaking address check fieldwork for resolution, and this evidence is collated to produce the final main address register. Recent address additions will be captured for inclusion in a supplementary print run.

<sup>8</sup> Lower Super Output Area (LSOA) - A geographic hierarchy designed to improve the reporting of small area statistics built from groups of Output Areas (typically 4 to 6) and constrained by the boundaries of the wards

Average household size analysis using patient register data as an estimate of the number of people <u>and</u> households by LSOA	x	✓
A comparison of the electoral register data provided by LAs with the electoral register data held by ONS	✓	x
A review of the comments received from LAs on the quality templates provided for the key comparator data sources	x	✓
A review of the supplementary data / research provided by LAs	✓	✓
Analysis of HMO data	✓	✓

The following sections give an overview of the data provided, analysis undertaken and key findings for each of the analyses listed above.

## 4. Council tax versus address register analysis

An important quality assurance check is the ability to validate the number of occupied and unoccupied households in the census against comparator sources. The council tax versus address register analysis allowed ONS to compare the number of households in the council tax data with numbers on the address register and identify any large differences. ONS was able to use the exemption and discount information to help investigate these anomalies. This analysis allowed ONS to assess the comparability of the council tax data with the address register to assess its value as a comparator source. It also enabled ONS to assess the value of the exemption and discount information.

<b>Criteria for LA inclusion in the analysis</b>	All LAs who provided council tax data including exemption and discount information
<b>Number of LAs included in the analysis</b>	Pilot LAs – 10, QA studies LAs – 23
<b>Data used</b>	Council tax data including exemption and discount codes – Provided by LAs  Census Residential Address Register – held by ONS
<b>Geographic level analysis available</b>	Postcode or above depending on level at which council tax data provided

### 4.1 Council tax data

#### 4.1.1 Information provided by LAs

At the QA pilot meeting in January 2010, the value of the council tax data, particularly the exemption and discount information, was widely discussed. A clear steer was provided about the importance of the source for understanding inconsistencies with census definitions. The council tax data is considered to be timely, is available at a low level of geography and is of a high quality. ONS already

---

used for 2001 Census outputs. LSOAs are of consistent size across the country, are not subject to boundary changes and have a minimum population of 1,000.

plans to use council tax data from the Valuation Office Agency (VOA). However, the VOA data does not include exemption and discount information.

There is potential to use council tax exemption and discount information in the supplementary quality assurance work and to QA second residences, single person households, student halls of residence and vacant properties. ONS currently has no source of council tax exemption and discount information as it is not collected centrally at the level of detail required. It was therefore agreed that pilot LAs would supply these data at LSOA level.

A variety of exemption and discount codes were provided and the data were supplied in a number of different formats. An important lesson learnt from the pilot was the need to be more prescriptive about the information required from the QA studies LAs. This included providing LAs with:

- a list of exemption and discount codes / descriptions (see Annex B)
- a specification of the preferred format of the data
- an outline of the metadata required

LAs were also asked to include any unbanded<sup>9</sup> addresses with their council tax data.

#### 4.1.2 The quality of council tax data

As part of the QA pilot, research work was undertaken to assess the quality of the council tax data for a subset of pilot local authorities. The main findings were:

1. Council tax data from LAs contained some invalid postcodes which are not recognised in the National Statistics Postcode Directory (NSPD). Results across several LAs showed a non-match average of 3 per cent
2. There were some postcodes outside the geographic boundaries of the LA
3. There were discrepancies between the definitions of a single dwelling. For example, in some cases a student hall of residence block is considered to be one exempt council tax unit (dwelling) while in other cases each flat or living space is recorded separately. These inconsistencies in recording were identified both within and across LAs

For the purposes of the pilot and QA studies analysis, postcodes were matched to the NSPD and also tested for validity. This involved testing for length, positioning of alpha and numeric characters, and alpha character combinations. The numbers of postcodes failing the checks or falling outside the LA boundaries were reported. However, as numbers were small, these postcodes were not removed from the analysis.

#### 4.2 Address register data

For the QA pilot, an extract of the draft address register was used referring to October 2009 (version 8A). This extract did not include communal establishments or feedback received from local authorities to help ONS resolve unmatched and partially matched addresses. It is recognised that the address register will be kept up-to-date or will be subject to improvements before its use in census but the draft address register provides an indication of likely household count at this stage.

Based on feedback from the pilot, for the QA studies the most up-to-date address register available at that time was used (version 12B<sup>10</sup>). Note this extract did not include communal establishments or new residential addresses identified from ONS address check fieldwork<sup>11</sup>.

---

<sup>9</sup> Unbanded property listings refer to addresses that have not yet had their council tax valuation.

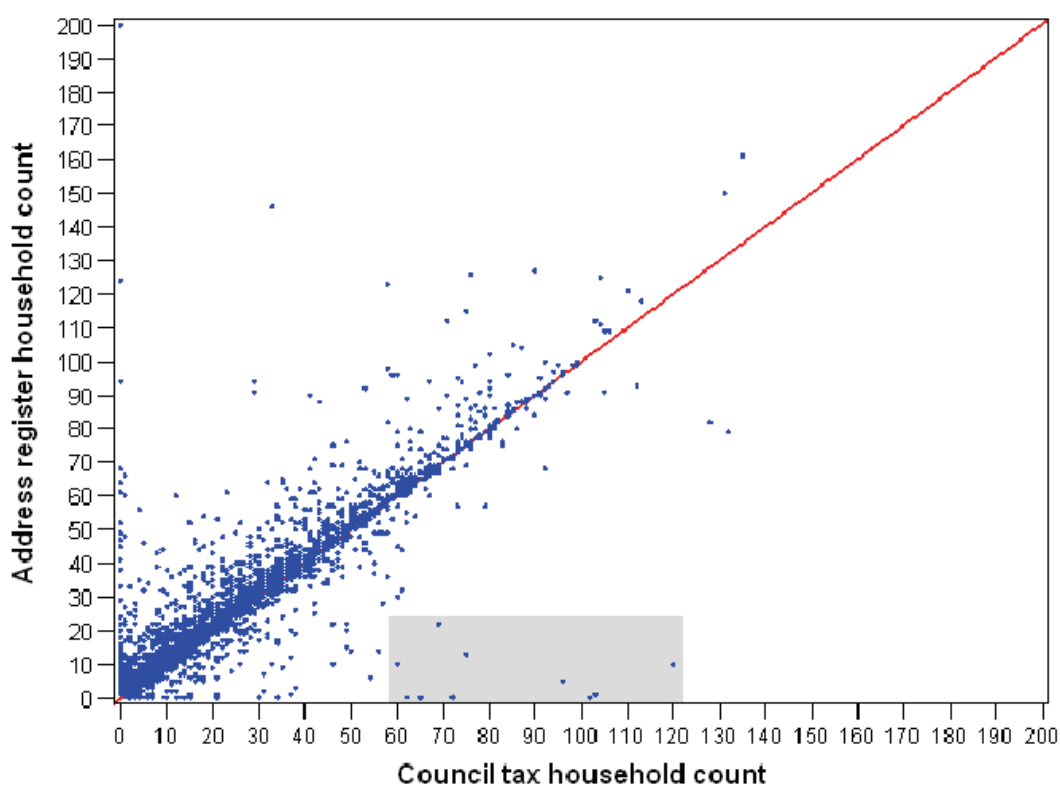
<sup>10</sup> Version 12B of the address register contains address source data up to May/June 2010 and is therefore considerably less up to date than the final 2011 Census Address Register will be. In addition, only one fifth of the over-coverage identified by the ONS fieldwork has been removed from version 12B (as the fieldwork was

### 4.3 Address register versus council tax analysis

A comparison of the council tax (CT) and address register (AR) data was made to assess the coverage of the two sources and better understand the definitional differences. Council tax exemption and discount code information was subsequently explored to assess whether it could explain inconsistencies.

Figure 2 shows a comparison of household counts at postcode level between the council tax and address register sources for an example LA. The x axis shows the number of council tax paying units from the LA supplied council tax data whereas the y axis shows the household count from the address register. Each point on the graph represents a postcode. Any points lying on the X=Y line are postcodes where the household count was the same on both sources.

**Figure 2 – Comparison of council tax and address register household counts for example LA**



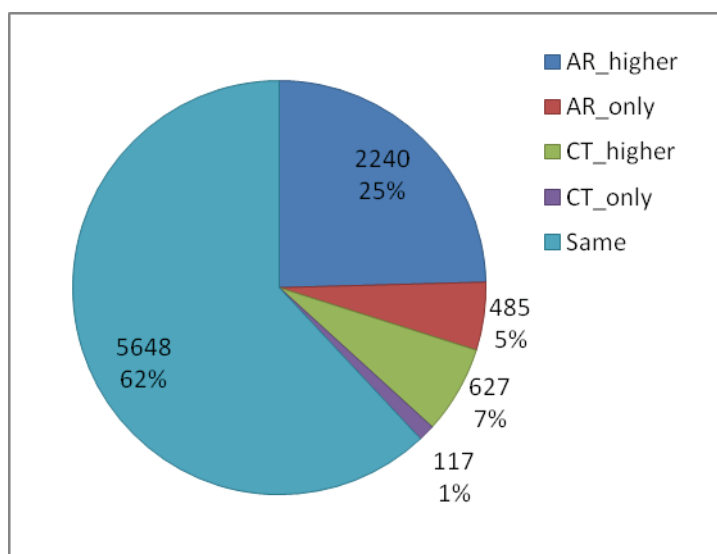
Once the postcodes were matched between the two data sources they were categorised based on the household count difference. For the example LA above, Figure 3 shows that 5,648 postcodes have matching household counts across the two data sources. This represents 62 per cent of the total postcode count. For postcodes where the household count is higher on the address register, Figure 3 shows that 2,240 or 25 per cent fall under this category.

---

not completed until late summer) and none of the under-coverage had been added to 12B (as new addresses found by the fieldwork will be included in the supplementary print run).

<sup>11</sup> ONS fieldwork has been undertaken in 15% of postcodes in England & Wales (postcodes selected using an area prioritisation model). All residential addresses on the address register within these postcodes have been checked to identify over-coverage (duplicates, addresses that no longer exist, commercial properties etc) and under-coverage (residential addresses within the postcode that are missing from the address register).

**Figure 3 – Postcode counts across matching categories**



Where large differences between the two sources were observed, a sample of postcodes was selected for further investigation (the grey area of Figure 2). The additional exploration focused on those postcodes where the council tax count was greater than the address register count. The council tax exemption and discount information was used to try to explain why these differences had arisen.

**Table 1 – Number of households within selected postcode outliers by council tax code**

Council tax code	Postcode A	Postcode B	Postcode C	Postcode D	Postcode E	Postcode F	Postcode G	Postcode H	Postcode I	Postcode J
<b>Full charge</b>	0	0	1	3	7	67	75	0	2	7
<b>Class C - Vacant</b>	0	0	0	1	0	1	0	0	0	4
<b>Class N - Students only</b>	19	33	63	90	113	0	0	103	100	0
<b>Class G - Occupation prohibited</b>	0	0	0	0	0	0	0	0	0	49
<b>Class M - Halls of residence</b>	43	39	1	1	0	0	0	0	0	0
<b>Class Q - Unoccupied</b>	0	0	0	1	0	0	0	0	0	0
<b>Class E - Unoccupied moved to hospital</b>	0	0	0	0	0	1	0	0	0	0
<b>Council tax total</b>	<b>62</b>	<b>72</b>	<b>65</b>	<b>96</b>	<b>120</b>	<b>69</b>	<b>75</b>	<b>103</b>	<b>102</b>	<b>60</b>
<b>Address register total</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>5</b>	<b>10</b>	<b>22</b>	<b>13</b>	<b>1</b>	<b>0</b>	<b>10</b>
<b>Difference</b>	<b>-62</b>	<b>-72</b>	<b>-65</b>	<b>-91</b>	<b>-110</b>	<b>-47</b>	<b>-62</b>	<b>-102</b>	<b>-102</b>	<b>-50</b>

Table 1 shows the postcodes that were selected for further analysis (shaded grey on Figure 2). For postcode A, the council tax data contains 62 more households than the address register. From the

council tax exemptions information, these properties are student only households and halls of residence. Part of this difference can be explained as the address register used for the QA studies work excludes communal establishments.

The figures of particular concern are those where full charge properties appear on the council tax data but not on the address register (Postcodes F and G).

For instances where differences cannot be explained, internet-based evidence has been used to explore areas which have experienced change, and to get a general feel for an area. This has proved to be an effective way of identifying new development although distinguishing between commercial and residential properties has sometimes been inconclusive.

#### 4.4 Council tax versus address register conclusions

The council tax exemptions and discount information is a valuable source of data for census quality assurance. It can be used in the QA of second residences, single person households, student halls of residence and vacant properties. It will also be used within supplementary quality assurance work to help understand anomalies and differences between census estimates of households and council tax totals.

An additional benefit identified by the QA studies work is the value of the council tax exemption data in recognising UK and foreign armed forces via Class O and P exemption codes. This information will be invaluable when checking estimates for LAs with armed forces.

Through work undertaken on the pilot LAs there were some data quality issues identified with the council tax data held by LAs. There may need to be some de-duplication and cleaning of the postcode data prior to its use in census QA.

Given how useful these data will be, ONS will be requesting council tax data with exemption and discount codes from all LAs in England and Wales at the time of the census. We will be asking QA studies and pilot LAs who have already supplied council tax data to provide an updated extract.

We will need data to be provided on a consistent basis, in a standard format, with appropriate metadata. Section 13 contains details of the information that will be requested from LAs.

## **5. Average household size analysis**

Based on feedback from the pilot LAs, ONS wanted to explore the value of using average household size as part of census quality assurance. Average household size brings together person and household counts to calculate the average number of people per household. Erroneous or extreme average household size figures will highlight potential problems with the census data for further investigation. In particular, this analysis will help to identify cases where:

- communal establishments have been incorrectly reported as households and vice versa
- continuation questionnaires<sup>12</sup> have not been requested and/or submitted resulting in an under-recording of large households
- there has been underestimation within households

---

<sup>12</sup> Continuation questionnaire – The questionnaire to be requested and completed by households with more than six household members or more than three visitors

This analysis allowed ONS to explore the potential of using average household size calculations to highlight areas of concern. It also provided an opportunity to assess the suitability of using the patient register on its own or in combination with council tax data as a comparator source for this check.

<b>Criteria for LA inclusion in the analysis</b>	All LAs who provided council tax data
<b>Number of LAs included in the analysis</b>	Pilot LAs – 10, QA studies LAs – 23
<b>Data used</b>	Council tax data including exemption and discount codes – Provided by LAs  Patient register data – held by ONS
<b>Geographic level analysis available</b>	LSOA level but analysis also undertaken at ward and postcode sector for LAs unable to provide LSOA

### 5.1 Council tax data

As described in section 4 above, local authorities were asked to supply their council tax data. From this information, we were able to derive the number of households by LSOA.

### 5.2 Patient register data

ONS has access to record level patient register data for England and Wales. At the time of the QA pilot, the only datasets ONS had access to were as at mid-year. These datasets excluded names and addresses but included the patient's home postcode. For the QA pilot, ONS used the patient register extract as at end June 2009 to derive the number of people registered with a doctor resident within each LSOA.

In May 2010, ONS received its first extract of the enhanced patient register. This extract includes names and addresses and refers to end March 2010. For the QA studies analysis, ONS was able to calculate the number of people registered with a doctor resident within each LSOA as in the QA pilot. However, the enhanced version of the patient register was also used to calculate the number of households. It is noted that where no one was registered as living at an address with a GP the household was missed.

Of the records in the patient register March 2010 enhanced extract (post de-duplication) for the QA studies areas, ONS was able to assign local authority codes to 99.4 per cent of the data using Matchcode (address management software used within ONS). The remaining 0.6 per cent of records, where there was insufficient address information to assign a geography, were excluded from the analysis.

### 5.3 Average household size analysis

Using the census population estimates, ONS will be able to calculate average household size based on individual and household returns. Figures that are extreme or implausible indicate a potential error in either the person and/or household count. ONS wanted to use administrative data to i) pilot this approach and ii) explore how these data can be used as a comparator to indicate when census distributions are outside expected ranges.



ONS calculated the average household size at LSOA level for all LAs where sufficient data was available. The analysis was used to identify outliers and implausible values at LSOA level which ONS then tried to investigate and explain. There were two calculations made:

1. ONS calculated the average household size by LSOA using patient register (PR) data as the numerator and council tax (CT) data as the denominator.

$$Average\_Household\_Size = \frac{Persons_{pr}}{Households_{ct}}$$

2. For the QA studies areas, ONS was able to calculate the number of households using the enhanced patient register data. For these LAs, the average household size was also calculated using the patient register data as both the numerator and the denominator.

$$Average\_Household\_Size = \frac{Persons_{pr}}{Households_{pr}}$$

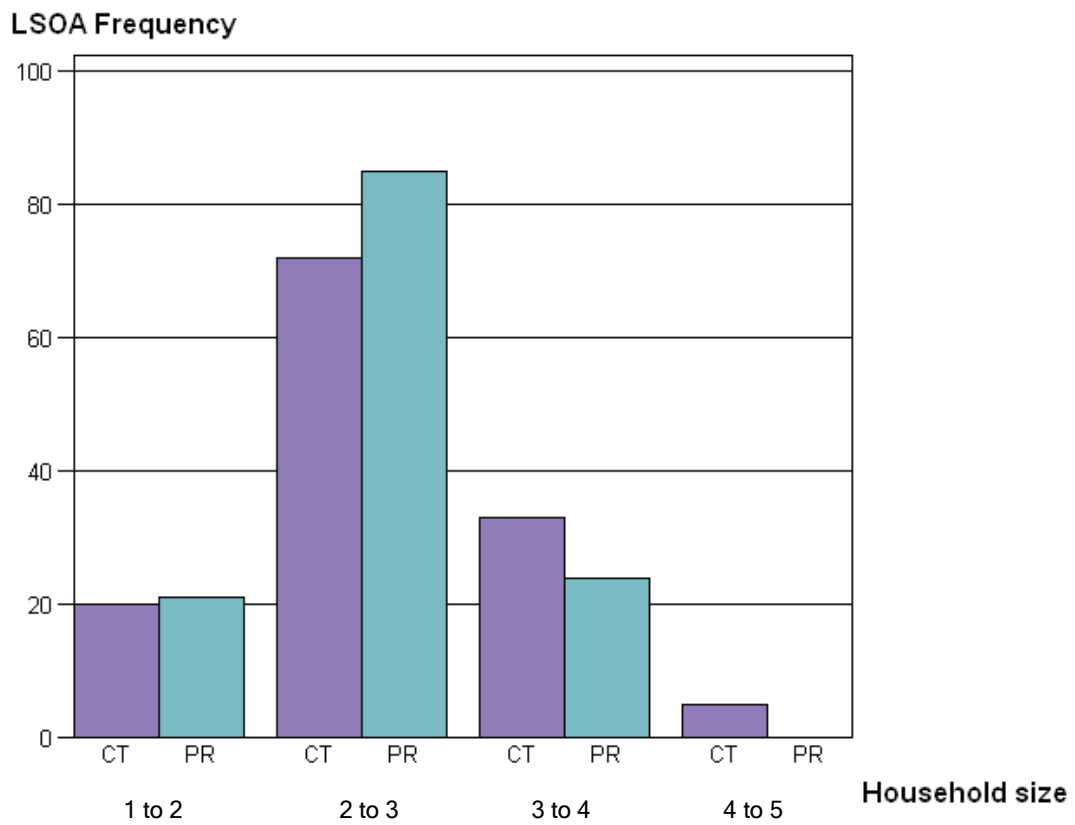
Figure 4 shows a typical analysis for one of the LAs participating in the QA studies. Each bar represents the number of LSOAs by average household size band. It compares figures calculated using council tax data as the denominator with rates calculated using solely patient register data.

Figure 5 compares the average household sizes directly based on the two calculations for the same LA. The x axis shows the average household size using the council tax data as a denominator whereas the y axis shows the average household size using the patient register data as a denominator. Each point on the graph represents an LSOA. Any points lying on the X=Y line are LSOAs where the average household size was the same using both methods.

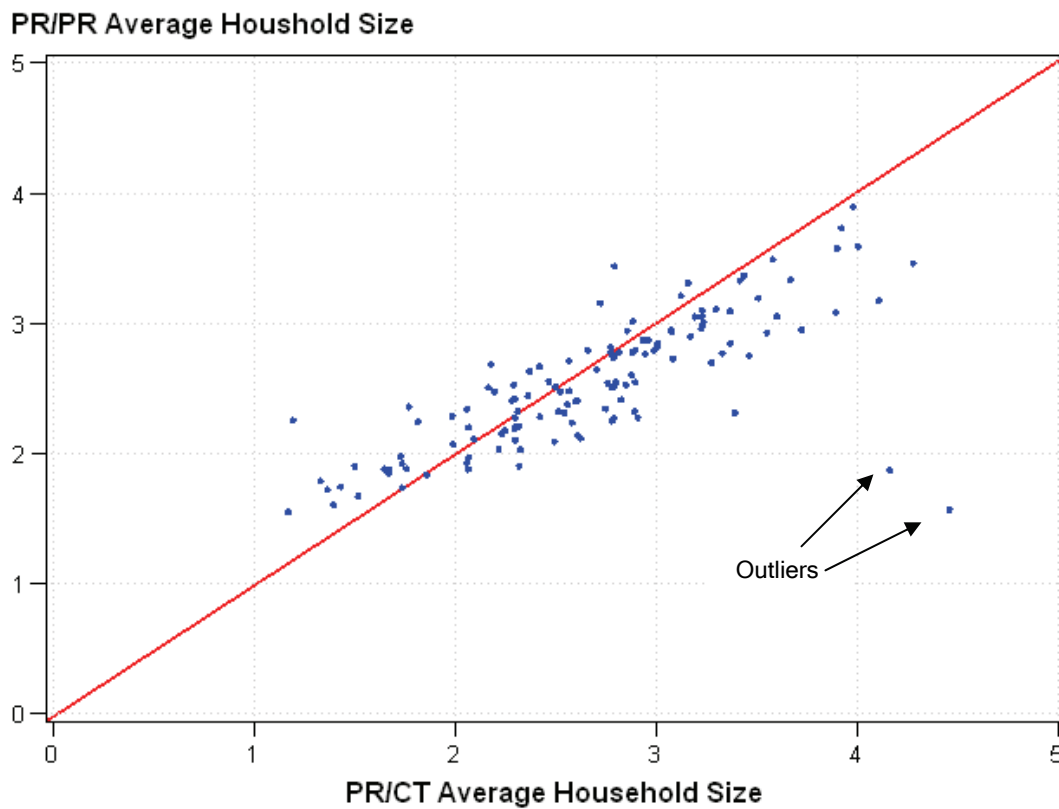
Table 2 provides additional information on the key summary statistics for the LSOAs within the LA.

For the example shown, the average household sizes using the two methods are broadly similar and within the range expected. There were 2 outliers where average household size using council tax data was much higher than using the patient register. These LSOAs were investigated further. Using the council tax exemption information, it was confirmed that both LSOAs contain a high proportion of households classed as university halls of residence. It is possible that these halls were categorised as a smaller number of households in the patient register data based on the addresses supplied.

**Figure 4 – Number of LSOAs within LA by household size bandings**



**Figure 5 – Comparison of average household sizes using two different calculations**



**Table 2 – Summary statistics for household size calculations**

	<b>Patient register/patient register household size (LSOA)</b>	<b>Patient register/council tax household size (LSOA)</b>
<b>Min</b>	<b>1.55</b>	<b>1.16</b>
<b>Max</b>	<b>3.90</b>	<b>4.46</b>
<b>Range</b>	<b>2.35</b>	<b>3.29</b>
<b>Mean</b>	<b>2.54</b>	<b>2.68</b>
<b>Median</b>	<b>2.51</b>	<b>2.73</b>

#### 5.4 Average household size conclusions

The conclusion of the analysis was that there is value in including an average household size check within census quality assurance. This check will flag implausible and extreme census average household size values for further investigation. It will also compare average household size, calculated using census data against values derived from administrative sources.

It is currently proposed that census values be compared against a diagnostic range that has been calculated using i) patient register data ii) data from the Integrated Household Survey and iii) Gemserv data<sup>13</sup>. Census values which fall outside this range may be explored further by exploiting council tax exemption and discount information.

## **6. Electoral roll analysis**

<b>Criteria for LA inclusion in the analysis</b>	Pilot LAs only All LAs who provided electoral registration data
<b>Number of LAs included in the analysis</b>	Pilot LAs – 7
<b>Data used</b>	Electoral registration data – Provided by LAs Electoral registration data – Held by ONS
<b>Geographic level analysis available</b>	Postcode or above depending on level at which electoral registration data provided by LA

### 6.1 Electoral registration data

A number of limitations with electoral registration data were discussed at the QA pilot meeting in January 2010. However, the source has potential for exploring anomalies within the supplementary QA work. Electoral registration data might provide information on country of origin, large households and occupancy.

<sup>13</sup> Gemserv data – Domestic energy consumption data (including counts of domestic meter points) provided by the Department for Energy and Climate Change.

At the QA pilot meeting in January 2010, local authorities were asked to provide a copy of their electoral register by LSOA (or lower level of geography), by country of origin and by age (17-18, 18+) as at December 2009.

ONS has access to the full register under the 'Representation of the People (England and Wales) Regulations'. It receives versions of the register which are published in December from the October canvass. For the purposes of this exercise the latest version of the ONS-held register was used. This was based on the October 2009 canvass.

## 6.2 Electoral register analysis

As part of the QA studies work, ONS wanted to compare the version of the register held by local authorities with the version held by ONS. The aim of the work was to assess the importance of timeliness on the register and determine whether electoral registration data should be separately requested from local authorities or whether information held by ONS can be used.

Local authorities provided electoral registration data by number of households, number of people or both. Data were provided either at postcode or LSOA level.

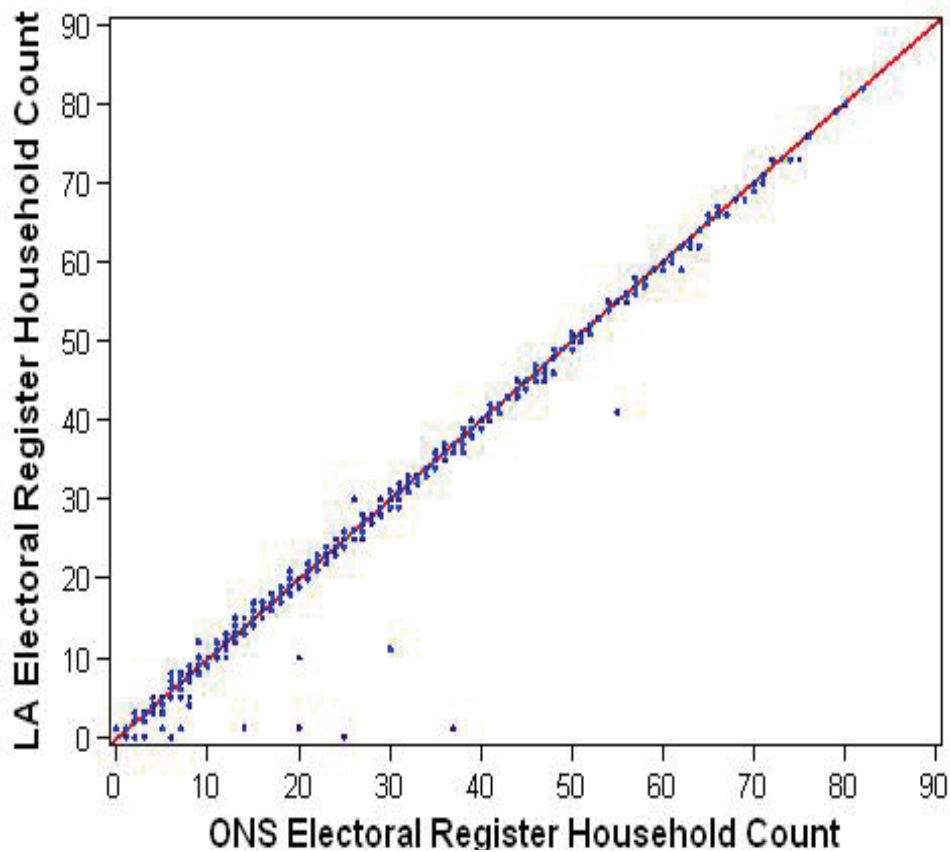
Data were analysed at the level of geography provided by the local authority. Counts of households by postcode/LSOA on the datasets provided by LAs were compared with the equivalent information from the register held by ONS.

Figure 6 below shows the analysis for one of the two local authorities able to provide their electoral registration data by counts of household at postcode level. The x axis shows the household count from the ONS-held register whereas the y axis shows the household count from the LA supplied register. Each point on the graph represents a postcode. Any points lying on the  $X=Y$  line are postcodes where the household count was the same on both registers.

Note that when analysing data down to postcode level, data quality issues are more noticeable than at higher levels of geography. We would expect to see variation in postcode level analysis due to the small numbers involved.

Figure 6 shows that for this LA the relationship between household counts on the two sources of data is very strong. In fact, 99% of the postcodes show a difference of less than 2 households between the ONS and LA data. A similar profile was observed for the other LA able to supply household information.

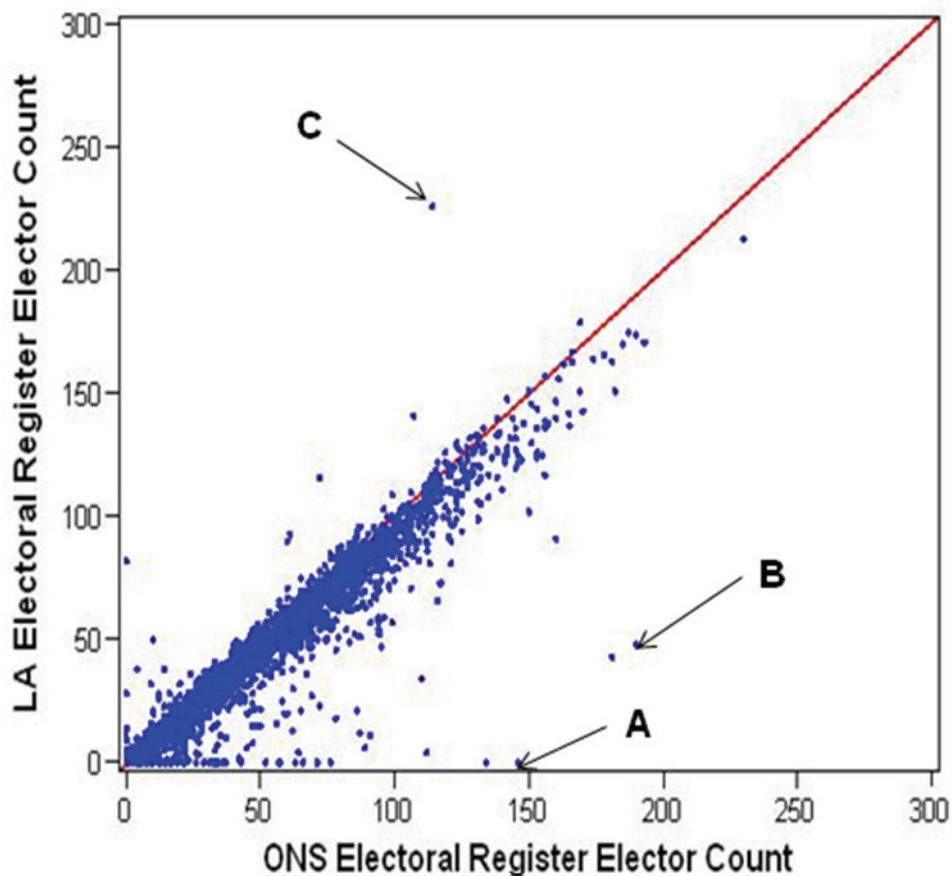
**Figure 6 – Electoral registration data for counts of households by postcode**



ONS undertook a similar analysis for those LAs able to supply electoral registration data by person count. Again a very strong relationship between the two sources for the majority of LAs was observed. However, Figure 7 shows the analysis for one local authority where the relationship was weaker.

ONS wanted to explore the postcodes where the difference between the two sources was greatest. The three postcodes which ONS investigated further are labelled A, B and C on Figure 7. Where data allowed, ONS extracted additional information about these postcodes from the electoral registration data and used internet-based evidence to help explain potential reasons for the difference. The findings from this further investigation are shown in Table 3.

**Figure 7 – Electoral registration data for counts of electors by postcode**



**Table 3 – Investigation into electoral registration analysis outliers**

POSTCODE	ELECTORS LA ER	ELECTORS ONS ER	DIFFERENCE	FINDINGS FROM INVESTIGATIONS
A	0	146	-146	Block of flats. Appears to be new development.
B	48	190	-142	Extra addresses. Appears to be new development.
C	226	114	112	Same block of flats as Postcode A. Potential discrepancies in postcodes between the two sources.

Further exploration work seemed to suggest that differences between the registers were due to new developments and a difference in timing of register extracts. This explanation was confirmed at the feedback session with the pilot LAs. Local authorities were able to confirm that the LA above had been going through a period of regeneration with significant new development. The analysis for those local authorities which had not experienced this level of new build showed a strong relationship between the ONS and LA registers.

### 6.3 Electoral register analysis conclusions

It was concluded that the relationship between the LA electoral registration data and the information held by ONS is generally very strong. For the purposes of census quality assurance, it was agreed that ONS should use the data it holds and not make a request for LAs to resupply the data specifically for QA work. However, it was noted that as in the case of the LA shown above, timeliness can be an issue for those local authorities experiencing a significant level of new development. For those local authorities, any supplementary evidence that can be supplied to advise on timing of the electoral register information will be welcome.

In light of the conclusions above, the electoral registration analysis was not repeated for the wider group of QA studies areas.

## **7. HMO analysis**

Houses in Multiple Occupation (HMO) data held by LAs was identified as a potential source of address intelligence to help ONS determine numbers and ensure these properties are covered by the address register. The HMO data is being used to target extra work at an early stage in the enumeration process and more generally within supplementary QA.

Further to the first meeting of the pilot LAs in January 2010, LAs were able to provide HMO data. This information was shared with the Address Register Team which, for a subset of LAs, was able to compare the lists provided with field activities to assess coverage.

The information received from the pilot LAs proved to be so valuable that the Address Register Team initiated a request for HMO data from every local authority.

## **8. Housing benefit data**

At the pilot meeting in January 2010 there was some discussion about the usefulness of housing benefit data for census quality assurance. Although this is a timely source of data, it only relates to people on low incomes in rented accommodation. Pilot local authorities felt housing benefit data is of insufficient quality and adds no extra value beyond the council tax information. They therefore did not recommend it as a comparator source for census quality assurance.

Aggregate housing benefit data was provided by two of the pilot local authorities. Analysis of these datasets was not progressed. For these two LAs, the housing benefit data will form part of the LAs' package of supporting evidence and if appropriate will be used within the supplementary QA.

Housing benefit data was not specifically requested from the QA studies LAs.

## **9. Quality templates**

As described in Section 2, ONS will be making extensive use of administrative and survey data sources in the census QA process. It is important for ONS to have a good understanding of each source being used in order to:

- assess its suitability as a comparator source
- plan when the data will be used and how it will be accessed
- identify definitional and timing differences to the census

- help determine what the range around the comparator source should be, outside which census data for a particular area will be flagged for further investigation

The team has completed a template for each of the comparator sources drawing on a range of evidence. The templates record the key features of the dataset and will evolve over time as knowledge and experience of the source increases.

The set of templates for the key comparators will help share knowledge of administrative sources across the team and will increase the QA panel's understanding of the datasets being used to accept / reject census estimates.

At the meeting on 7 July, the QA studies LAs had the opportunity to review the quality templates for council tax, Higher Education Statistics Agency (HESA), patient register and School Census data and feedback their comments. They have subsequently also reviewed quality templates for birth registrations, Defence Analytical Services Agency data on armed forces, DWP Migrant Worker Scan data and CLG Household Projections.

LAs were asked to draw on their knowledge and practical experience of these datasets to provide feedback and general observations on the quality of these sources. They were encouraged to identify issues and problems with the datasets specific to their areas, providing evidence-based analysis / reports where possible. They were also asked to suggest additional headings and improvements to the template itself.

The LAs returned very useful information on the eight quality templates provided. These comments and suggestions have been used to update and improve the templates supplied. They are also being used in the development of the checks and will be available to refer to during the QA process.

A copy of the quality template is included in Annex C.

## 10. Supplementary evidence

Local authorities in the QA pilot and QA studies were encouraged to provide ONS with any supplementary evidence, including:

- survey and administrative sources (not already held or used by ONS) that ONS will consider as comparator data sources for the census
- feedback on local issues, concerns, expectations and observations that might impact on the estimated population of the LA, for example the prevalence of particular population sub-groups
- local feedback on significant issues that might impact on census quality
- any relevant research and analysis, including the documentation, findings and recommendations from commissioned research

A diverse range of evidence was received which ONS has reviewed and is currently considering how best to use, not just within the Census Quality Team but across the census operation.

ONS will be compiling a pack of supplementary evidence for each LA using this information which can be referenced during the QA checking process to provide additional insight and local evidence when understanding and interpreting checks.

### 10.1 Summary of supplementary evidence received from LAs



Local authorities provided a combination of reports and data. A summary of the information received is listed below.

#### 10.1.1 Reports

- Local intelligence reports
- Mid-year estimates of population: Response to ONS migration statistics consultation
- Reports relating to representations on the 2001 census
- University managed accommodation: Issues with enumeration
- Caravan sites: Issues with enumeration
- Local Land and Property Gazetteer (LLPG) linkage information
- Independent Schools Survey 08/09: Details of LA survey
- Residential Development Surveys
- BME survey
- Strategic housing market assessment (SHMA)
- Counting confirmed and unconfirmed population – Mayhew Harper Ltd. report
- Detailed population analyses based on administrative sources
- Student term-time address analysis
- Somali Community information
- Language and ethnicity analyses and reports
- CE lists
- Aggregate counts of Flag 4s by nationality – Trends by country

#### 10.1.2 Data

- University accommodation
- Empty dwellings / vacancy rates
- Dwelling completions
- Social housing by LSOA
- Bus passes by LSOA
- Libraries data

### 10.2 Example of supplementary evidence which will provide valuable insight during QA process

One local authority was able to provide borough level aggregate counts of the number of Flag 4s by country of birth. Flag 4s are codes within the Patient Register Data System which indicates an individual was previously resident outside the UK.

As this level of detail is not available within the patient register data held by ONS, information such as this is a useful source of supplementary evidence for the census QA work.

The information provided clearly shows how migration patterns have changed over recent years and how the proportion of Eastern European migrants has increased substantially compared to the number of international migrants within this LA.

This evidence will be useful when exploring anomalies in the key demographic checks and to help understand migration patterns in the area.

Given the value of this information, ONS will be requesting these data from all local authorities. ONS recognises that country of birth is not routinely collected in all areas as part of the patient registration process so some local authorities will be unable to provide these data.

## 11. Lessons learnt

The QA studies project provided the team with a valuable opportunity to understand the datasets, familiarise themselves with the statistical tools and explore different analytical methods. The key outcomes are listed below.

- Improved understanding within ONS of the data sources held by LAs
- Legal position of requesting and holding data from LAs clarified
- Secure data transfer and storage options refined
- An early insight into sources and checks ONS had already planned to use
- Opportunity to improve data analysis skills
- Data handling procedures improved
- Analytical methods, in particular the use of scatter plots to compare datasets and identify outliers, incorporated into the Automated QA system
- The value of using internet-based evidence confirmed for areas that have experienced housing change
- The importance of definitional and timing issues when comparing datasets recognised
- Potential refinements to original plans identified. This includes the use of an average household size check and the value of electoral registration data as a comparator source
- The value of council tax data including exemptions and discount codes confirmed
- Confirmation that the diagnostic range approach should be taken forward and refined
- National collection of HMO data initiated
- Useful feedback on the quality issues of key comparator datasets received
- Importance of engagement with LAs for knowledge-sharing and better understanding of data sources recognised

## 12. Recommendations

1. Request for all local authorities to provide council tax data including exemption and discount information to be issued
2. Patient register data by country of birth to be requested from all local authorities
3. ONS to review information submitted through the Census Local Partnership Plans to identify other sources of supporting evidence which will be useful for census quality assurance
4. Council tax exemption and discount data to be used as part of supplementary QA to investigate cases where census household estimates fall outside expected ranges, and to reconcile specific population sub-groups, for example UK and foreign armed forces
5. Check on Average Household Size to be included as part of the quality assurance process
6. Electoral registration data held by ONS to be used as part of the quality assurance process

## 13. Continuing engagement with LAs

### 13.1 Request for additional data

Further to the recommendations above, ONS is writing to local authority chief executives and assistant census liaison managers in December 2010 to request data and information that will be useful for the QA process. The information being requested is summarised in sections 13.1.1, 13.1.2 and 13.1.3.

Each LA should record that they have provided this evidence to support the QA process in their Census Local Partnership Plan, along with a brief description of the nature of the information supplied.

#### 13.1.1 Council tax exemption and discount data

ONS requests that all local authorities provide their council tax data (including exemption and discount information) to ONS. The letter being sent to all LAs will outline the information required. This will include a specification of the preferred data content, data format and information regarding the legal basis on which the data will be supplied.

ONS would like these data to be provided as at census day (27 March 2011) or as close to that period as possible. Data are required as soon as feasible after census day but preferably prior to early June 2011.

#### 13.1.2 Patient registration data by country of birth

ONS requests that, where available, local authorities provide the number of new patient registrations and new immigrants by country of birth at local authority level for the year up to end March 2011.

ONS recognises that country of birth is not routinely collected in all areas as part of the patient registration process so some local authorities will be unable to provide these data. It should also be noted that ONS already holds record level patient register data for England & Wales. It is specifically the country of birth information, which is not part of the ONS patient register extract, that is requested from LAs.

Data are required as soon as feasible after census day but preferably prior to early June 2011.

#### 13.1.3 Supplementary evidence

Local authorities in the QA studies have provided a wide range of supplementary evidence which can be referenced during the QA process. Local authorities not included in the QA studies and those in the QA studies with additional information to share are encouraged to submit any relevant materials to ONS. It would be helpful for this information to be supplied prior to the census where available but no later than early June 2011.

### 13.2 Quality Assurance Advisory Group

Building on the effective partnerships that have developed during the QA studies between ONS and local authorities, ONS is keen to continue to draw on the expertise and local knowledge of users to help in the development of the quality assurance processes. It has established a Quality Assurance Advisory Group, consisting of a small number of the local authorities which were highly engaged in the QA studies project.

The objectives of the advisory groups are:

- to keep local authorities involved and up to date with current thinking as the QA development work progresses
- to make best use of the expertise of local authorities and obtain a local authority perspective on a) the methods, b) the QA processes and c) how best to engage with the wider local authority community
- to provide local authorities with confidence in the processes being followed in ensuring the census estimates are fit for purpose

The advisory group includes representatives from the following:

Cambridge City Council  
Ceredigion Council  
Greater London Authority  
The Royal Borough of Kensington and Chelsea  
Manchester City Council  
Newcastle City Council  
Newham London Borough Council  
Nottingham City Council  
Oxford City Council  
Westminster City Council

## **ANNEX A - Data sources that will be used in the QA process**

- ONS mid-year population estimates
- 2001 Census data
- Birth, death, marriage and civil partnership registers
- Electoral register
- Patient register
- DWP Lifetime Labour Market Database
- DWP pensions data
- DWP Migrant Workers Scan
- HMRC Child Benefit data
- DfE and WAG School Census data
- Independent School Census data
- Council tax data
- CLG Houses in Multiple Occupation
- CLG Gypsy Traveller Accommodation Assessment
- DASA armed forces data
- MoJ prisons data
- HESA data
- DfE data on boarding schools
- CLG National Register of Social Housing

### Selection of admin sources that are being investigated

- DWP WPLS and CIS
- ONS Inter-Departmental Business Register (IDBR)
- GEMSERV – Multiple electricity meters at an address
- CLG housing data (for example sales data, building statistics, household projections)
- WAG housing sources
- Data for other communal establishments, (for example care homes, hospital and hostel data)
- DVLA data
- Data on caravan sites/parks

## **ANNEX B - Council tax discount, exemption and liability codes and descriptions**

### Discounts

S	Single Person Discount
DO	Disregards Only – solely occupied by persons disregarded for council tax reasons
SD	Single Disregard – occupied by just one person who is disregarded for council tax purposes
SWD	Single with Disregard – all but one adult in the dwelling are disregarded
SH	Second Home discount
SE	Standard Empty
D	Disabled Relief Award

### Exemptions

**Class A** Vacant dwellings where major repair works or structural alterations are required, under way or recently completed (up to twelve months).

**Class B** Unoccupied dwellings owned by a charity (up to six months).

**Class C** A vacant dwelling (i.e. empty and substantially unfurnished) (up to six months).

**Class D** A dwelling left unoccupied by people who are in prison.

**Class E** An unoccupied dwelling which was previously the sole or main residence of a person who has moved into a hospital or care home.

**Class F** Dwellings left empty by deceased persons.

**Class G** An unoccupied dwelling where the occupation is prohibited by law.

**Class H** Unoccupied clergy dwellings.

**Class I** An unoccupied dwelling which was previously the sole or main residence of a person who is the owner or tenant and has moved to receive personal care.

**Class J** An unoccupied dwelling which was previously the sole or main residence of a person who is the owner or tenant and who has moved in order to provide personal care to another person.

**Class K** An unoccupied dwelling where the owner is a student who last lived in the dwelling as their main home.

**Class L** An unoccupied dwelling which has been taken into possession by a mortgage lender.

**Class M** A hall of residence provided predominantly for the accommodation of students.

**Class N** A dwelling which is occupied only by students, the foreign spouses of students, or school and college leavers.

**Class O** Armed forces' accommodation. (Included as exempt dwellings rather than chargeable dwellings because contributions in lieu will be paid in respect of them).

**Class P** A dwelling where at least one person who would otherwise be liable has a relevant association with a Visiting Force.

**Class Q** An unoccupied dwelling where the person who would otherwise be liable is a trustee in bankruptcy.

**Class R** Empty caravan pitches and boat moorings.

**Class S** A dwelling occupied only by a person, or persons, aged under 18.

**Class T** A dwelling which forms part of a single property which includes another dwelling and may not be let separately from that dwelling, without a breach of planning control.

**Class U** A dwelling occupied only by a person, or persons, who is, or are, severely mentally impaired who 'would otherwise be liable to pay the council tax, or only by one or more severely mentally impaired persons and one or more students, students' foreign spouses and school and college leavers.

**Class V** A dwelling in which at least one person who would otherwise be liable is a diplomat.

**Class W** A dwelling which forms part of a single property, including at least one or other dwelling, and which is the sole or main residence of a dependent relative of a person who is resident in the other dwelling.

Owner (landlord) liability flags – where liability is retained by owner of property who does not live there

A	Care Homes, nursing homes and some hostels
B	Dwellings of religious communities
C	Houses in multiple occupation
D	Residences of domestic servants
E	Residences of ministers of religion
F	Asylum seekers

## **ANNEX C - Copy of quality template**

**DRAFT Quality Template – insert name of data source**

<b>Overview</b>	BRIEF overview  i.e. summarise what it is, who collects it etc. Should be no more than a short paragraph
<b>Dataset title</b>	
<b>Location</b>	Where is the data held? E.g. is it in the Virtual Microdata Laboratory (VML)? What is the drive and folder if a team member needs to get access?
<b>Classification</b>	Is it an administrative dataset or collected for statistical purposes e.g. survey?
<b>Data owner/supplier</b>	Who owns the data and which department supplies it to ONS?
<b>Access to the data</b>	<ul style="list-style-type: none"> <li>• Does ONS already hold the data?</li> <li>• Is there a legal gateway established (or soon to be established)?</li> <li>• Are there any specific access issues e.g. limited number of people in the team can access School Census data?</li> </ul>
<b>Prior knowledge</b>	<ul style="list-style-type: none"> <li>• How is this data source used by other areas of ONS? e.g. Patient register used to validate MYEs and derive internal migration stats  e.g. Improving Migration Statistics Programme (IMPS) matching School Census to MYEs</li> <li>• Which area in ONS has used this data and has significant knowledge of the data source?</li> </ul>
<b>Lowest output area</b>	<ul style="list-style-type: none"> <li>• Is the data available record level or aggregate?</li> </ul>
<b>Data collection</b>	<ul style="list-style-type: none"> <li>• Who is the original source of the data? e.g. HESA collected by universities</li> <li>• How it is collected?</li> <li>• What is the reason for data collection?</li> </ul>
<b>Data cleaning</b>	<ul style="list-style-type: none"> <li>• What cleaning is carried out on the data before it is received by ONS? e.g. are duplicates in School Census data from schools and pupil referral units (PRUs) removed before we receive it?</li> </ul>
<b>Variables</b>	<ul style="list-style-type: none"> <li>• Include a complete set of variables collected as bullet points. Split into sections to make clear which are mandatory variables and which are voluntary and which variables correspond to census ones</li> <li>• Explain where there are definitional differences to census in the variables collected, e.g. School Census ethnicity classification is different to census</li> </ul>
<b>Intended use</b>	<ul style="list-style-type: none"> <li>• Provide details of the specific checks that ONS intends to carry out using this data source as a comparator</li> </ul>
<b>Geographic coverage</b>	<ul style="list-style-type: none"> <li>• Does the data source cover UK, England and Wales? England only? Local authorities? Is the geography compatible with census?</li> </ul>
<b>Topic coverage</b>	<ul style="list-style-type: none"> <li>• Which population sub-groups are included? e.g. Does it cover people, households and communal establishments (CEs)?</li> </ul>



	<ul style="list-style-type: none"> <li>Which groups are excluded? Does it exclude a sub-group of people e.g. only include people of a certain age or with certain characteristics?</li> </ul>
<b>Data quality</b>	<ul style="list-style-type: none"> <li>Start with a section on national issues, including general positive aspects e.g. reasons why a source makes a good comparator and general negative issues</li> <li>Include sub-headings for each data quality issue to make it clear and readable. Where evidence has been provided in relation to one of the data quality issues, e.g. from an LA, provide a brief explanation and provide a separate link to the detailed evidence / research report (if applicable).</li> </ul> <p><i>This is the section that would expect to build up the most over time</i></p>
<b>Timeliness</b>	<ul style="list-style-type: none"> <li>Are time series data available? For which years?</li> <li>How often (and when) are the data updated, collected and received by ONS?</li> <li>When are data available to census?</li> <li>What are the latest data available?</li> <li>As at census day, what is the time reference point of the data available</li> </ul>
<b>Concepts and definitional differences with census</b>	<ul style="list-style-type: none"> <li>Include specific information on how definitions for key variables relate to the census e.g. definition of student in HESA compared to census</li> <li>How does the population covered by the source at a national level relate to the census definition of usual resident of the UK i.e. that census only includes people who have stayed or intend to stay in the UK for 12 months or more</li> <li>How does the population covered by the source at a sub-national level relate to the census definitions of usual residence within the UK i.e. that census counts people with more than one address at their permanent or family home</li> </ul>
<b>Anecdotal issues raised by LAs</b>	<ul style="list-style-type: none"> <li>Include bulleted summary of the points raised at the QA studies meeting. Once these are substantiated with evidence etc they can be moved into the main report</li> </ul>
<b>Summary of usefulness of source</b>	<ul style="list-style-type: none"> <li>Include bullet points of key reasons why we intend to use the source in QA e.g. high level of coverage? Timely data? Accurate source?</li> </ul>
<b>Work schedule</b>	<ul style="list-style-type: none"> <li>Summary of future work being carried out in relation to the source</li> </ul>
<b>Related links</b>	<ul style="list-style-type: none"> <li>Provide links to full reports e.g. administrative sources quality reports, other material produced by ONS that evaluates the source</li> </ul>
<b>History of template development</b>	<ul style="list-style-type: none"> <li>When created and updated (and by whom)?</li> <li>Who has reviewed the contents?</li> <li>Has it been sent to LAs?</li> </ul>
<b>Proposed</b>	<ul style="list-style-type: none"> <li>List of possible people to review template e.g. relevant topic lead or person</li> </ul>

<b>reviewers</b>	
------------------	--