3

The 2011 Census

From preparation to publication

Garnett Compton, Amy Wilson and Brian French

3.1 Introduction

The planning and execution of the decennial population census is a vast and complex undertaking often described as the largest peacetime operation carried out in the UK. It is certainly the largest statistical exercise, aiming to collect socio-demographic information from every individual and household in the four home nations. Census-taking is a complex exercise because it involves the deployment of thousands of staff at peak times, the co-ordination of services across a number of different suppliers, the management of people and teams remotely, engagement with a wide range of stakeholders with a wide variety of interests and, crucially, connecting with every person and household in the country. The most recent census was undertaken in 2011 and was one of the most successful censuses of modern times.

Census-taking is a devolved responsibility in the UK. The Office for National Statistics (ONS) is responsible for the census in England and Wales, whereas in Scotland responsibility lies with the National Records of Scotland (NRS) and in Northern Ireland with the Northern Ireland Statistics and Research Agency (NISRA). The ONS is the department responsible for UK outputs and delivering UK statistics to Eurostat as part of the European Council (Framework) Census Regulation (EC No. 763/2008) (Eurostat, 2008). As a result of devolved responsibilities, the design and operation of the three censuses will differ in some aspects to meet local requirements and/or operational needs. However, a key aim of each of the Census Offices in the 2011 Census was to harmonise the census operation and most importantly the census outputs as far as possible. This was achieved through close liaison and co-operation between the three UK Census Offices. The National Statistician and the Registrars General (ONS, 2005) have a formal agreement to work together to achieve consistent and comparable census outputs – both to meet domestic users’ requirements and to fulfil the UK’s international obligations.

This chapter covers the main aspects involved in the taking of the 2011 Census, the vast majority of which are common between the three censuses. However, within the operation, there are areas where it was designed or implemented differently by the three UK Census Offices and, where these are significant to the story of how the 2011 Census was undertaken, these differences will be highlighted. The chapter seeks to give a flavour of some of the new approaches and methods that were introduced to ensure the delivery of a good-quality census that would meet
the needs of users and address the challenges of a changing society. Given those complexities, only the key areas are explored; a more in-depth explanation of the planning, implementation and evaluation of the 2011 Censuses is available in each of the General Reports for England and Wales (ONS, 2015a), Scotland (NRS, 2015) and Northern Ireland (NISRA, 2015).

The key areas explored in the next five sections of this chapter are as follows: Section 3.2 summarises the planning and preparation for the 2011 Census, including testing and user engagement; Section 3.3 outlines the development of the questionnaire; Section 3.4 explains the data collection operation and key supporting processes such as publicity and stakeholder engagement; the data processing operation is documented in Section 3.5 and the production and dissemination of outputs are documented in Section 3.6. The chapter ends with some conclusions in Section 3.7.

3.2 Planning and preparation for the 2011 Census

Recommendations from the 2001 Census and strategic aims

The design of the 2011 Census took into account the lessons learned from the 2001 Census, as assessed by the Census Offices’ own evaluations, and also the changes in society that were expected between 2001 and 2011. In addition, in England and Wales, recommendations made by external bodies, such as the Treasury Select Committee (House of Commons Treasury Committee, 2002), the National Audit Office (2002), the Public Accounts Committee (House of Commons Committee on Public Accounts, 2003) and the Local Government Association (2003) informed the design. For England and Wales, the key issues raised in these reviews covered the need to:

- select external suppliers of outsourced census operations early, using rigorous procurement procedures, and test their systems before the census;
- increase the efficiency of census questionnaire delivery by developing a high-quality and up-to-date address list;
- enable better central control of field processes and activities by developing robust field management and questionnaire-tracking systems;
- have earlier and more detailed engagement with stakeholders, particularly local authorities, and review consultation processes to ensure the disabled community’s needs were taken into account;
- ensure that the views of people in Wales and the Welsh Government are better reflected in census planning, by reviewing consultation processes;
- review whether or not the coverage survey’s design is sufficient to identify under-enumeration in the hardest-to-count areas;
- review the need to collect information on income;
- review the cost–benefit trade-offs in aiming to produce more timely outputs that are consistent and harmonised across the UK; and
- review the mechanisms to protect statistical confidentiality without eroding the utility of the data.

The separate internal reviews carried out in Scotland and Northern Ireland, whilst not so detailed, made recommendations in the same general areas.

Taking account of these and many other comments arising from the 2001 Census, the design of the 2011 Census in England and Wales was based on a number of broad strategic aims:
to give the highest priority to estimating correctly the national and local populations;
• to build effective partnerships with other organisations, particularly local authorities, in planning and executing the field operation;
• to provide high-quality, value-for-money, fit-for-purpose statistics that meet user needs, inspire user confidence, and are consistent, comparable and accessible across the UK as far as possible;
• to maximise overall response rates and minimise differences in response rates in specific areas and among particular population sub-groups; and
• to protect, and be seen to protect, confidential personal census information.

The strategic aims informing the census design for Scotland and Northern Ireland were similar to those for England and Wales.

**Key elements and innovations of the 2011 Census design**

To achieve these aims, and to respond to changes in society since 2001, the design of the 2011 Census was significantly different from its predecessors. The societal changes included: an increasingly ageing population; a more mobile population with more complex living arrangements; increasing numbers of migrant communities, particularly from eastern European countries; and greater numbers of people in both single-person households and in dwellings with multiple household occupation.

The key elements in the design were that:

• the census would aim to cover everyone usually resident in the UK on census night, with a subset of information also collected from visitors present in households on census night;
• questionnaires would primarily be delivered by post, using a purpose-built address register;
• field staff resources would be focused in areas which were particularly hard to enumerate, and from which initial response rates were low;
• the public could return completed questionnaires either by post or online;
• help would be available to anyone who had difficulty in completing the census questionnaire;
• there would be a slight increase in the number of questions compared with the 2001 Census, although the questionnaire would be re-designed to make it easier to complete;
• stringent confidentiality and security procedures would protect the information gathered in the census and would conform to the requirements of census confidentiality, data protection and freedom of information legislation, as well as to the provisions of the Statistics and Registration Service Act 2007;
• to help achieve the public co-operation that a census relies on, there would be publicity to convey the purpose and value of the census, and to give assurances about the confidentiality with which information is treated;
• initiatives would be put into place to maximise, and measure effectively, the quality of the information collected; in particular, census coverage and quality surveys would be carried out to measure the number of people not counted by the census and the quality of the responses given, quality assurance panels would review the outputs prior to publication to ensure differences with other sources were explained (thereby increasing users’ confidence in the estimates and realising the benefits of the census); and
• the statistical outputs from the census would be designed to meet user requirements, and dissemination would be to a timetable.

These initiatives and other key elements of the census design are shown in Figure 3.1.
3.3 Meeting the information needs of users: the census questionnaire

The investment of time and resource in a national census can only be justified if the results are accessible to users and meet their information needs. To determine these needs, in accordance with the principles and practices in the Code of Practice for Official Statistics (UKSA, 2009), the Census Offices consulted widely with users of census data. In order to understand user needs, formal public consultations were held, supported by a number of national open meetings about particular issues, including equality-related questions on ethnicity, identity, religion and language. As part of this, the views and needs of government departments, local authorities, the health service, the academic community, the business sector, voluntary sector and local communities were collected.
The topics and questions chosen for inclusion on the 2011 Census questionnaires were those with the most demonstrable need by users, and for which census questions could be devised to produce reliable and accurate data. In each case, no other comparable and accessible source of information was available that would enable multivariate analysis with other census topics. As in the 2001 Census, the demand for questions was much more than was feasible and/or appropriate in a single census questionnaire. This demand led to challenging decisions for the Census Offices in deciding on questions and balancing the need for change against the advantages of continuity. There was a slight increase in the number of questions compared with the 2001 Census, and a re-designed questionnaire to make it easier to complete. In particular, these were:

- new questions on national identity and citizenship;
- additional response categories in the ethnicity question, which differed between countries;
- new questions on second residences;
- a new question on language;
- the inclusion of a civil partnership category in the marital status question;
- questions on date of entry into the UK for immigrants and intention to stay;
- the omission of questions on access to toilet and bath/shower;
- in both Scotland and Northern Ireland there were questions on long-term health conditions and travel to place of study; and
- in Northern Ireland, new questions on adaptation of accommodation, Ulster-Scots and volunteering.

Developing and finalising harmonised questionnaires across the UK was particularly challenging given the devolved responsibilities for many policy and service delivery functions. Inevitably, there were some differences in the content and design of the three census questionnaires. The comparability of these questions and the subsequent results derived from the information collected are set out in the 2011 Census UK Comparability Report (ONS, 2015b).

**Developing the questionnaire**

A large programme of qualitative and quantitative question testing was undertaken to develop the questions. The testing consisted of five different activities: (i) cognitive testing, i.e. one-to-one interviews designed to ascertain whether or not a question was acceptable and worked as intended; (ii) a number of small-scale postal tests were conducted to collect sufficient quantitative information about questionnaire design and in particular the success of new questions; (iii) an opinions survey was used a number of times to test understanding and acceptability of new or changed questions (e.g. the migration questions) or towards colour terminology with the ethnicity question; (iv) a number of focus groups were held to understand the acceptability of particularly challenging topics and questions, including ethnicity, national identity and migration; and (v) large-scale tests in 2006 (Scotland) and in 2007 (in England, Wales and Northern Ireland) were used to test the acceptability, quality and impact on response through the inclusion of a question on income.

The near-final questionnaire was used in census rehearsals across the UK in 2009 and the final topics and questionnaire were endorsed by the respective legislatures, with the passing of the respective Census Orders and Census Regulations.
3.4 Data collection operation and key supporting processes

**Using an address list as the spine of the census operation**

As set out earlier, a key objective of the 2011 Census was to maximise overall response and minimise variation in response between local authorities. A critical element of this was the introduction of an up-to-date address list that in England, Wales and Northern Ireland would be crucial to the delivery of questionnaires – using Royal Mail to post questionnaires to the vast majority of households. Hand delivery of questionnaires by enumerators had been the design for each of the previous censuses. However, moving to a post-out methodology provided savings in enumerator time that could be focused more on supporting and encouraging responses from those that did not respond. It was also found that in the 2001 Census more than a third of households visited by the enumerator failed to make contact and ended up effectively posting a questionnaire anyway.

Questionnaire tracking and deployment of a flexible field force ended up effectively enabling the Census Offices to track each questionnaire through a unique identification number assigned to each dwelling at an address. This unique ID was then associated with the paper questionnaires as a barcode which was scanned when questionnaires were returned to the processing site. This enabled the Census Offices to have daily information on which addresses had not responded and prioritise field staff to areas where the response was lower than expected. This was a significant finding from the 2001 Census where there was insufficient information to accurately assess and manage the progress of the field enumeration.

Therefore, the design hinged around having an address list that had high levels of coverage to reduce the risk of under-coverage in the census operations. If households were excluded from the list, they were unlikely to receive a questionnaire and therefore had a higher propensity not to respond. However, the list also needed low over-coverage (such as commercial addresses misclassified as residential, duplication, demolished properties and addressing errors) as these would lead to wasted postage, wasted field staff time resolving these in the field and the risk of irritating households who had already completed and returned their census questionnaires. Although hand delivery of questionnaires was still used for the vast majority of addresses in Scotland, the printed questionnaires were personalised to the address. Therefore, the accuracy of the address register was still vitally important to ensure that the correct questionnaire was delivered to the correct household.

There was no single national source of addresses that provided sufficiently recognised quality. Therefore, the Census Offices developed a single source through matching the main national address source Royal Mail Postal Address file (PAF), the local government maintained National Land and Property Gazeteer (NLPG) and the Ordnance Survey/Ordnance Survey NI address layer products and on-the-ground verification of address anomalies. More information on the development of address lists is set out in the General Reports of the three Census Offices.

In evaluating the 2011 Census, it was estimated that the address registers had under-coverage rates of about 1 per cent and over-coverage rates of nearly 4 per cent. Using information from the Census Coverage Surveys (CCS), the Census Offices were able to estimate and adjust for those households who did not complete the census, including those who were not on the original address list. The over-coverage rates were similarly estimated and resolved through field staff visits to these addresses and the CCS. However, these did have an impact on the field force and the public, in particular when the over-coverage was due to duplication of addresses.
Managing and deploying field staff

The field operation was designed to maximise the likelihood of achieving the strategic aims of the census. The main field operation comprised four distinct activities: delivery, collection, follow-up of non-responders and special enumeration.

The delivery of questionnaires to households was described earlier with the vast majority of households in England, Wales and Northern Ireland receiving their questionnaires through the post and the vast majority in Scotland receiving theirs via hand delivery. However, communal establishments and managed residential accommodation were enumerated in a more traditional manner with special enumerators visiting each communal establishment to identify the type of establishment and the number of residents, leaving the appropriate number of questionnaires. The public could respond to the census in several ways: (i) completing the paper questionnaire and posting it back to the processing site in the prepaid envelope; (ii) completing an online questionnaire (each paper questionnaire had a unique internet access code that let the householder enter the census website securely and complete their questionnaire online; this code linked their online return to their address); and (iii) handing their completed paper questionnaire to census staff on the doorstep.

The use of post-out and post-back, and online completion, were very cost effective ways to enable households to make a return with minimal effort. As a result of these changes, the 2011 Census in England and Wales was able to halve the overall size of the field force employed from about 70,000 employed in 2001 to just 35,000. At the same time, the amount of resource put into follow-up was three times more than in 2001, enabling field resources to be targeted at those unwilling or unable to make a return without support and/or encouragement. In effect, the entire field operation was designed to focus effort and resources on non-responders by: changing the field staff roles (and training) to concentrate primarily on collection; and having flexible workloads so that field staff were not assigned to one specific area.

This shift in approach required a different field staff allocation model. Rather than assigning a collector to a particular area with the sole responsibility of collecting questionnaires from non-responding households in their area (previous census models), collectors were assigned larger areas but all had the same workload (similar numbers of non-responding households). To do this, a model was developed to estimate the amount of follow-up resource required to achieve the minimum response rate thresholds for each lower super output area (LSOA). Inputs to the model included: estimates of how many households would return a questionnaire without follow-up (the initial return rate); how successful a collector would be in securing a completed questionnaire from each visit during follow-up; and how long each visit would take given the type of area (its hard-to-count classification). These inputs were based on evidence from other social surveys, the 2009 rehearsal and the 2007 test.

Generally, the right amount of resource was allocated to each area. As with any model, there will be errors in the various input parameters and to account for this some caution was used in those areas where the model was expected to have more variability. There were compensating errors in various assumptions of the model. The initial return rate was higher than expected and more visits per hour were achieved. These were offset by follow-up success being much lower than expected. However, the flexible allocation of field staff and being able to move them around during the operation and focus more resource on areas with lower-than-expected response rates (using the daily information from the questionnaire tracking system) meant that the field operation met its overall targets of maximising response and that all local authorities exceeded the minimum levels of response.
Online collection

For the first time, the 2011 Census offered households and individuals the opportunity to complete their return online, as an alternative to the traditional paper questionnaire. About 16 per cent of returns were completed through the secure online census; the majority of completed questionnaires were returned by post, as had been the case in 2001. Although the level of online response was lower than expected, the online service was regarded as a success, providing a number of benefits to the data collection operation. It:

• met the expectations of both the public and census stakeholders for an online questionnaire;
• provided an environment in which the security of the census information could be better protected;
• improved overall responses by offering an alternative to householders who may have been less inclined to complete a paper questionnaire;
• delivered a more accessible census for the disabled community; and
• avoided the need to scan and capture a significant proportion of the returns, thereby speeding up, and reducing the cost of, data processing.

Most importantly, the online service was easy to use, improved data quality by prompting for missing responses, limited the scope for incorrect responses, did not fail at any time during the process and had no security breaches. Indeed, security was the highest priority requirement in developing the online system. Confidentiality of personal information is a cornerstone of the Census Offices’ assurance to the public, and any breach of data security would not be tolerated.

The daily volume of online returns for England and Wales from 21 March to 22 May is shown in Figure 3.2. Although these data are for England and Wales, the patterns of response are

![Figure 3.2 Daily volume of online census returns, England and Wales, 2011](source: ONS (2012f, Figure 1, p. 2).)


consistent across the UK. As expected, census day (27 March) had the highest volume of daily returns with about 19 per cent of the total online household returns. Apart from census day, the daily volumes of returns throughout each week peaked on a Monday.

Other notable influences on the volume of returns included: the increase from Saturday 19 March to census day following the change of media message to ‘Fill it in now’; an increase on Wednesday 6 April, which coincided with the start of collectors visiting the homes of non-responders (there was a rise on Tuesday 26 April – after the Easter weekend – which probably reflected people returning from the weekend and finding reminder cards left by a collector); and rises during the week of 2 May, which reflected the last week of collectors’ visits before the start of non-compliance. The great majority of respondents took between 10 and 20 minutes to complete the online questionnaire. This varied of course depending on the number of people in the household. The average time taken to complete the questionnaire was 22 minutes.

The experience of running an online questionnaire, the first time by UK statistics offices, will be invaluable for informing the 2021 Census and other data collection operations. More information about the timings, location and characteristics of those that did respond online are available in the ONS report on ‘Providing the Online Census’ (ONS, 2012a).

**Online help and support**

In addition to the online completion facility, an online help system was available in English and Welsh, and Gaelic in Scotland. The system provided general information about the census, information about census questions, and answers to questions that the public might have. Answers to specific questions such as who to include on the questionnaire, how to obtain replacements, additional questionnaires or large-print versions of questions and translation leaflets proved very popular. The Help website received nearly 5 million visits. The system was especially effective at helping people decide how to complete certain parts of the questionnaire. The online help was also a popular method for requesting materials such as audio cassettes, Braille booklets or British Sign Language DVDs.

The census help-line handled over a million telephone calls during the operation. A large number of queries related to address register ‘anomalies’, most of which related to: houses that had been split into flats and had received only one questionnaire; houses that used to be flats and were no longer sub-divided which had received multiple questionnaires; properties that were incorrectly addressed (such as ‘Basement flat’ instead of ‘Lower flat’, or ‘Flat A’ instead of ‘Flat 1’); and Welsh road names that were more commonly known in their English form, and *vice versa*.

**Publicising the 2011 Census**

The success of the census depended on making contact with every UK household. This presented a unique challenge to marketing communications as it necessitated engaging with every household in the UK and motivating them to fill in their census questionnaire. The increasing diversity of households made such contact difficult for key population groups such as ethnic minorities, migrants and young adults – which are some of the very groups for which census information is critical. It was impossible to reach all census audiences through one communications channel. Therefore, various communications channels and activities, including paid-for, owned and earned (free) media channels, were employed depending on the target audience.
The paid-for media channels for the campaign were: TV advertising; outdoor/out-of-home advertising; digital and social media; Black and minority ethnic (BME) radio and print advertising; and magazine advertising targeting young adults. Pro-active media relations, for example news releases, worked alongside advertising and paid-for communications to generate free coverage on radio, TV and print.

**Local authority and community engagement – key partnerships**

A census – which encompasses the whole population – has an exceptionally large number of stakeholders with varying degrees of influence and interest. For the 2011 Census, the UK Census Offices took a more systematic and consistent approach to communicating with stakeholders. In particular, extensive programmes of local authority and community liaison were initiated that were appropriate and relevant to each of the Census Offices. In particular, the Census Offices developed programmes that were aimed at working with local authorities and community representatives to utilise their knowledge and understanding of their local areas and communities.

To help support the strategic aims of the 2011 Census, the Census Offices initiated a programme of local authority liaison and engagement with the aims of: raising awareness and understanding of the 2011 Census; explaining the role that local authorities could play in participating and supporting the census; and building trust and confidence in census methodology.

Previous censuses have shown that certain population groups are less likely to complete and return their questionnaires. These include young men, certain black and minority ethnic groups, the very elderly, low-income families, non-English speakers and disabled people. The reasons for low response vary from an unwillingness to complete the questionnaire (because of concerns or misunderstandings of how the information is used), through to potential barriers such as a lack of English language skills. Community liaison activities were therefore designed and established to work with community representatives to:

- identify the main groups at risk of not responding and understand barriers to participation;
- initiate contacts with groups to communicate the purposes of the census and how data from it can be used;
- identify resources to interact with all community planning partnerships across the UK on census-related issues; and
- implement a system to identify ‘champions’ to promote the census from the largest identifiable group.

Some of the specific activities that flowed from these partnerships were:

- access to current, established, and effective networks and communication channels set up by such communities;
- access to translators (for example, contact with Northern Ireland Centre for Ethnic Minorities);
- designing tailored approaches to what best suits different groups (for the traveller community, for example, invitation to present and interact at a traveller-related event organised by Belfast City Council);
- partnership with influential and strategic organisations, including the Royal National Institute of Blind People (RNIB) and the Royal National Institute of the Deaf (RNID), that
resulted in targeted information campaigns and sign-posting to help-lines and facilities provided; and
• hints and tips on best methods of engagement.

Engagement with these local authority and community partners started before the rehearsal in 2009 and continued through the publication of outputs.

Confidentiality, security and privacy

The Census Offices recognised that the public needs to be confident that personal information collected in the census will be held securely. As in previous censuses, assurances were given to the public that all the information provided would be treated in the strictest confidence. An independent information assurance review (ONS, 2011a) was carried out prior to and during the census operation, covering a wide range of planning, management and implementation activities. The review team noted that, from the outset, ensuring the protection of personal information provided by the public had been a core objective in planning for the 2011 Census. They concluded: “As a result of our review, we are very satisfied that the three Census Offices are managing Information Assurance pragmatically, appropriately and cost-effectively. We are, therefore, confident that they are capable of delivering their IA [information assurance] objectives and that information will be handled in line with best practice and Government standards. The public can be assured that the information they provide to the 2011 Census will be well protected” (ONS, 2011a, p. 5).

The information collected in the 2011 Census was used solely for the production of statistics and statistical research. The Census Offices applied a statistical disclosure control process that modified some of the data before the statistics were released. The method employed was record swapping, which always introduces some uncertainty as to whether the value of any given small count is the true value. These measures proved satisfactory for protecting statistical confidentiality within the published census outputs.

3.5 Processing the results and data quality

In all, more than 27 million census questionnaires required data processing, which began by scanning the paper questionnaires and automatically capturing their data which were then merged with data captured through the online questionnaires. The data were validated to ensure that the values for each question were within the range specified in the relevant coding frame and that there were no duplicate responses. Coders assigned numerical values to written text and ticked boxes, applying coding rules and standardised national coding frames, such as SIC07 (Standard Industrial Classification 2007) and SOC2010 (Standard Occupational Classification 2010). This processing was carried out at specially commissioned and secure sites in Trafford Park, Manchester, for England, Wales and Northern Ireland returns, and in Livingston for Scottish returns.

After capture and coding, the data were passed to the UK Census Offices for further processing to validate, adjust and prepare for producing and disseminating results. This involved a number of key stages that are outlined in Table 3.1.

Overall, the methods and data sources used to capture, clean, validate and quality assure the census results were transparent and gave users confidence in the process and hence the census results.
<table>
<thead>
<tr>
<th>Processing activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range checks</td>
<td>This process checked that the value of each variable was within the valid range for that variable. For example, there were four valid values for the sex variable: male, female, missing or multi-tick. The 'range checks' process verified that all values for the sex variable were one of these four valid values. If an invalid value was found it was set to 'missing'. Missing values were then imputed as part of the edit and imputation process.</td>
</tr>
</tbody>
</table>
| Removing false person records | A person record was created during the recognition phase each time at least one mark was detected in any of the person questions. Records could be created in error if, for example, there was dust on the scanners that was incorrectly interpreted as a mark. This process identified genuine person records and those found not to be genuine were flagged as an invalid person record. For a person record to be counted as a genuine response and kept in the data, the following information had to be present on the record:  
  - name (from individual questions or household members table) or date of birth; and  
  - at least one other item, different from the above filter, from: name (from individual questions), date of birth, sex, marital status, or name (from household members table). |
| Resolving multiple responses | This process resolved both household and individual multiple responses at the same address. There was an increased likelihood of multiple responses from the same household occurring in the 2011 Census because of the introduction of online completion. A multiple household response was created if a paper and an online response were returned for the same address. A multiple individual response was created if a person was included on the same questionnaire more than once. Multiple responses were identified by looking for more than one response for an address ID or by looking at all of the individuals within a household for the same individual more than once. Name, date of birth, and sex were the variables used for the primary matching. When multiple responses relating to the same household or individual were identified, the records were merged to leave just one record for the household or individual. The most complete response was kept, with any missing variables being filled in from the other response(s) if possible. In the case of multiple individual responses, a response on an individual questionnaire was given priority over a response on a household questionnaire. Across the UK, the numbers identified and removed as multiple responses were relatively low. For example, in England and Wales, 237,000 individual records (0.4 per cent) and 181,000 household records (0.8 per cent) were identified and removed as being multiple responses. More information on these early processes is set out in the ONS report, 'Data Capture, Cleaning and Coding' (ONS, 2012b). |
| Edit and imputation | Respondents to any census sometimes make mistakes in their answers. This results in missing data or invalid responses which are inconsistent with other values on the questionnaire. An edit and imputation method was used to correct inconsistencies and estimate missing data whilst preserving the relationships between census characteristics (ONS, 2012c). |
The coverage assessment and adjustment operation helped the Census Offices to adjust for the number of people and households not counted in the 2011 Census. The extent of this undercounting was identified using the post-enumeration Census Coverage Survey (CCS covered approximately 400,000 households in the UK). Standard statistical estimation techniques were then used to produce an adjusted database from which the final census results were produced. These results also formed the new 2011 base for the mid-year population estimates in each country. More information on coverage assessment and adjustment is detailed in Chapter 4 of this book and in ONS (2012d).

In Northern Ireland, the CCS was supplemented by the Census Under Enumeration (CUE) Project. The CUE Project was initiated to augment the coverage of the census enumeration (that is, completed questionnaires) by using activity-based administrative data from the medical system to supply core information on non-responding households. Activities such as the collection of a prescription, changes to registration details and treatment by a dentist or optician were considered to provide good evidence of residence. The administrative data were considered an additional source of information along with the CCS. The CUE methodology was piloted and refined during both the 2007 test and the 2009 rehearsal.

The results of the CUE/CCS were then matched, at the individual level, to the corresponding 2011 Census data, identifying the number and characteristics of those missed in the census. The combined census and CCS information, along with statistical models, were used to produce an estimate of the number of people missed by the census. The people and households estimated to have been missed were then added to the database using similar techniques and processes to the edit and imputation stage.

In approximate terms, 94 per cent of households were enumerated through completed questionnaires, 4 per cent through the CUE and 2 per cent through the CCS. More information can be found in the Northern Ireland General Report of the 2011 Census (NISRA, 2015).

It is expected that this approach will be further developed by the UK Census Offices, particularly given the anticipated increased use of administrative data to augment and quality assure the 2021 Census.

QA procedures were built into all stages of data processing, and the 2011 Census estimates were subject to a rigorous QA process prior to their release (see ONS, 2011b). The overall aim was to provide confidence in the estimates by using comparator datasets and by conducting a series of vital checks.

The Census Offices carried out their own QA, but adopted the same general approach which is detailed in Chapter 4. Key steps in the process were as follows:

- a range of quality assurance panels reviewed estimates at varying levels of detail, including different geographic levels;
- a range of evidence was considered, including comparison with administrative data sources;
- the quality assurance process checked persons and their key characteristics (for example, students, armed forces, ethnicity);
- estimates of households occupied by usual residents were also quality assured;
- identifying issues which were adjusted for in the data processing; and
- further analysis to explain inconsistencies with the comparator data against which census estimates were evaluated.
Table 3.1 (Continued)

<table>
<thead>
<tr>
<th>Processing activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigning output geographies</td>
<td>This activity assigns each person and household record a number of geographies based on the address information collected in the census (such as usual residence, workplace address, second address). These are then used to allocate the records to any particular output geography, such as output area or workplace zone.</td>
</tr>
<tr>
<td>Statistical disclosure control</td>
<td>This activity applies the statistical disclosure control measure, record swapping, to protect confidentiality and minimise the risk of accidentally disclosing information about an individual or household in the main census outputs (see ONS, 2012e).</td>
</tr>
<tr>
<td>Creating derived variables</td>
<td>Some of the census results use variables derived from more than one census question; for example, age is derived from date of birth, and distance travelled to work is estimated from the location of the addresses of the place of usual residence and the place of work. This activity seeks to create all of the required derived variables.</td>
</tr>
</tbody>
</table>

3.6 Outputs

The 2011 Census provides the most complex and comprehensive set of information about the population ever produced. The recent growth in the demand for information, especially through the internet and social media, has encouraged a high user expectation regarding content and delivery that includes collaboration and user participation. The ultimate benefits of the census are only realised when the users of census data make use of the published outputs. Therefore, the investment of time and resources in a census can be justified only if the results are made accessible and the outputs produced meet users’ needs. Other chapters in Part V of this book provide excellent examples of the wide range of research uses that census results have been put to.

The Census Offices used their experience of the 2001 Census, user feedback and newly developed technologies to maximise census benefits, delivering a suite of products and services which included:

- the web as the primary dissemination route, with minimum paper products;
- increased ability to analyse census data online, for example through data explorer functionality and interactive graphics;
- comprehensive metadata delivered alongside the data;
- utilities to enable bulk download of census results via the web;
- larger set of products, ensuring maximum analytical use;
- microdata products (provided via secure mechanisms as appropriate);
- the provision of updated data for inclusion in the three UK Longitudinal Studies;
- licensed access for trusted users to more complex outputs that do not satisfy disclosure control requirements for public availability; and
- the provision of outputs to meet the requirements of EU regulations.

From the first release of 2001 Census results onwards, the Census Offices sought feedback on all aspects of census outputs. User forums such as the long-running Census Advisory Groups for England and Wales and Northern Ireland, and the Population and Migration Statistics Committee for Scotland, and specific interest groups such as the Microdata Working Group, provided the Census Offices with mechanisms to understand ongoing and evolving needs. The 2011 Census
outputs teams also carried out focused consultations using mechanisms such as web surveys, blogs and wikis. The outcomes from all these consultations were collated and used to influence the decisions taken in defining the 2011 Census output programmes.

As far as possible, a UK-wide approach was taken to understanding users’ high-level output requirements, with ONS, NRS and NISRA collaborating on different aspects of the user consultation programme on planned output. Users from all sectors – academic, commercial, central government, local government, health and others – confirmed that the planned design of outputs reflected their interests in, and needs for, both the retention of comparability with 2001 and information on new topics to be collected in 2011.

The main outcomes from the consultations that influenced the design and development of the census outputs were:

- for comparability and continuity purposes, only minor or no changes were made to the majority of existing 2001 Census outputs, other than those necessary due to changes in questions, or for statistical disclosure control purposes;
- to preserve stability, some outputs had additional age breakdowns incorporated and the statistical disclosure control method of targeted record swapping was developed to address the additivity and consistency problems arising from the post-tabular application of small cell adjustment in 2001;
- comparability documentation was published, highlighting areas of change between 2001 and 2011 and describing how outputs may be affected;
- the outputs reflect the width and breadth of the data collected on those topics included for the first time; and
- the generation, for the first time, of alternative population bases was a particularly popular innovation, including the bases for short-term migrants, workplace, workday/daytime, and out-of-term populations.

**Output geography**

Geography is a key element of census outputs. Every statistic produced from the 2011 Census is available for at least one of the various administrative or statistical geographies in the UK. An overall aim for the 2011 Census was to provide outputs in line with the National Statistics (NS) geography policy (ONS, 2015). The policy sets out the principles for using geographic information to produce and disseminate statistics. Its principles are driven by the objectives to:

- reference statistical events accurately, consistently and at as low a level of geographical referencing as possible;
- maximise the comparability of ‘National Statistics’;
- minimise the impact of changing area boundaries on National Statistics outputs; and
- provide the framework for defining and standardising how geographies and associated information are defined, used and presented in the production of statistics.

The 2011 Census results for output geographies are aggregations of whole output areas, which have been best-fitted to the higher geographies that were current at 31 December 2011. This is the method used to produce all 2011 Census and National Statistics, so that statistics produced on the same geography are consistent, comparable and non-disclosive. The exceptions to this are the exact-fit estimates for local authority areas (to which whole super output areas (SOAs) and thus
output areas (OAs) are constrained), workplace zones (which have been created by merging or splitting whole OAs (Ralphs, 2011) and national parks (because best-fit estimates were considered to be inappropriate for this largely rural geography).

In Northern Ireland, the main outputs were based on a fully-nested hierarchical geographic structure of Small Areas (based largely on 2001 OAs) within SOAs within Electoral Wards within Local Government Districts. The main outputs were based on the administrative geography in place in 2011, namely the local government boundaries in place since a review held in 1992.

**Dissemination**

The vision for 2011 was that the web would be the primary dissemination route, and would offer users easy navigation and functionality to customise outputs, charting and thematic mapping. To achieve this, the Census Offices’ existing web services were enhanced. Both Northern Ireland and Scotland provided enhanced access and functionality to the 2011 outputs through their websites and a Web Data Access programme was initiated to provide new functionality on the ONS website for the dissemination of England and Wales, and UK results. These channels were the primary vehicle for the publishing of the census standard products.

Technological developments have enabled the Census Offices to present data in more innovative ways, including infographics and data visualisations. Infographics, in particular, are an effective way to summarise census data and highlight key insights as indicated in Chapter 10 of this book. Similarly, data visualisations allow users to explore different variables and have more control over what they want to see. As a result, 2011 Census data are more relevant to a wider audience, as evidenced by the large numbers viewing the data and using the online tools.

The Census Offices strove to meet users’ requirements for statistics at varying levels of detail, for a number of geographies, subject to the overriding requirement to protect statistical confidentiality. The main products disseminated through the Census Offices’ websites include univariate and multivariate statistics, supported by statistical bulletins, short stories and data visualisations, including infographics and interactive maps. The key products made available from the Census Offices include standard and specialist census products. The former include:

- **Key Statistics** – largely percentages of selected key variables, designed to enable easy comparison across the geographies for which they were produced (for example local authorities).
- **Quick Statistics** – provided more detail on the breakdowns, or classifications, within a single census topic or variable, for output areas and higher (for example, a full breakdown of the ethnic group categories or single year of age population for a given geography).
- **Local Characteristics** – a series of more complex datasets designed to be available at output areas and all higher geographies. Often contained cross-classifications of more than one topic, such as age, sex, general health and ethnicity.
- **Detailed Characteristics** – similar to the local characteristics, but provided more detail within the topics (for example, full ethnic group classification) and available at a higher geography – mostly at middle layer super output area (MSOA) level and above.

As in previous censuses, the Census Offices have produced more specialist products aimed at specific audiences, such as the expert user. These include:

- **Small population groups** – subject to disclosure control restrictions, the detailed characteristics of some small population sub-groups, such as the Ravidassian and Nepalese communities who had made strong representations to ONS prior to the census. The threshold for table
production is 50 or more qualifying people in any given MSOA. Separate sets of outputs have been developed for areas where there are 100 or more, and 200 or more, people from the same small population group.

- **Microdata files** – (often referred to as Samples of Anonymised Records (SARs)) have been produced from each census since 1991. These datasets comprise files containing a sample of individuals or households drawn from the census database that have been anonymised and made available through different channels. Available are: (i) a public 1 per cent sample of persons; (ii) a safeguarded file, with a sample size of about 5 per cent, through the UK Data Service (UKDS) and after application, accessible via users’ desktops; and (iii) two secure files available through the ONS’ virtual microdata laboratory (VML) and made available only to approved researchers. Containing maximum sample sizes of 10 per cent, these are similar to the 2001 Controlled Access Microdata Samples (CAMS) for both households and individuals.

- **Origin–destination data** – (also known as flow data) comprised the travel-to-work and migration patterns of individuals, cross-tabulated by key variables of interest (for example, occupation). New products for the 2011 Census, however, provided the flow patterns separately for those living at a student address one year before the census, and also provided data on the movement of people between their usual address and any second address on which information was collected for the first time. A large number of the origin–destination outputs are at UK level, providing flows for usual residents within and between England, Wales, Scotland and Northern Ireland. However, for the 2011 Census, the UK statistical disclosure control policy required that the disclosure protection of the most detailed origin–destination tables should be controlled, in the main, through access only via ONS’ secure environment, the VML. This is a change from the 2001 Census, where the protection for similar outputs came from the post-tabular small cell adjustment that still allowed wide and easy access, but which also adversely affected the utility of the outputs. There are, however, a small number of less detailed tables that are available publicly.

- **Alternative population bases** – The main output base for the 2011 Census results is usual residents. Some basic demographic outputs using population bases other than usual residents are available from the 2011 Census. These use information from a combination of different census questions (such as workplace address) to focus on alternative population bases, including:
  - workplace population: figures for a given geography during standard working hours, taking account of the number of people who, for example, travel into a city to work (effectively a geographical redistribution of the usually resident population who are in work, allocated to their place of work);
  - workday population: figures for persons present in any given geography during the day, including non-residents with a workplace in the area, but excluding residents with a workplace outside the area;
  - out-of-term population: figures for a given geography, including students counted at their non-term-time address (that may or may not be the same as their term-time address).

**UK-based statistics**

In addition to the information produced and disseminated by each of the three Census Offices, the same statistics for the UK as a whole were produced (such as those required to fulfil international obligations as well as meet domestic users’ requirements):

- population figures by single year of age and sex; for the UK and all local authorities (or equivalent) in the UK, along with UK historic population pyramids (1911 to 2011);
• Key Statistics and Quick Statistics for all local authorities;
• Key Statistics and Quick Statistics for OAs, SOAs and equivalents in the UK; and
• EU outputs – a set of 60 multivariate tables consistent across the UK to meet UK requirements under EC regulation 763/2008. This information is available through the Eurostat Census Hub and provides comparable information across all member states.

These statistics, together with interactive maps and tools, are available on the 2011 UK Census web pages.

Making the 2011 Census data accessible was paramount. However, widening the census user base is possible only when potential users are educated about how they can benefit from the available data; potential new users also require help on how to access the data. Recognising the importance of these activities in maximising the benefits of the census, the Census Offices implemented a number of activities to promote and enhance the use of census data.

Case studies are a good and easy way to showcase different uses and benefits of census data to potential new users. The chapters in Part V of the book provide evidence of how census data are used by the academic research community. The Census Offices used their websites to illustrate how people/organisations can benefit from the 2011 Census. These pages show examples of the many ways of using census data and how different organisations from the private, public and voluntary sectors also benefit. Furthermore, toolkits, factsheets and instructions to help get people started with different census products were also made available to download. Some examples of uses and promotion outside the academic sector include:

• The Northern Ireland Housing Executive (NIHE) commissioned research to develop a census-based model of future housing need in Northern Ireland. In addition, on behalf of the NIHE, 2011 Census data have been used to model three key housing quality indicators at local government district level in relation to strategies on fuel poverty, decent homes and dwelling unfitness. This research, which benefited from additional work by Census Office staff in relation to the full 2011 Census dataset, enabled a reduction in the sample size of a recent Northern Ireland House Condition Survey and realised cost savings in the region of £250,000.

• In Scotland, NRS has supported a number of training programmes for prospective and practising school teachers and helped them develop practical skills in accessing and interpreting the statistics. The training has been delivered to trainee teachers in mathematics and social subjects through Continuous Professional Development, and at a range of conferences including the Scottish Mathematical Council, the Scottish Association of Geography Teachers and the Royal Statistical Society.

• As part of the Curriculum for Excellence in Scotland, all teachers have a responsibility for developing numeracy, both within the early stages of a broad general education and within the senior phase of school education. In addition, the Curriculum for Excellence emphasis on developing learners’ capacities as ‘responsible citizens’ and ‘effective contributors’ lends itself well to accessing and interpreting statistics about society. Studying Scotland’s people is already a feature of Education Scotland’s support materials but there was a clear opportunity for NRS to more widely publicise the use of the census outputs in schools and communities.

• Bristol City Council works with its partners through the Children and Young People’s Trust to deliver a wide range of services. From education, health and welfare to youth services and play, these services support the 92,000 children and young people in Bristol and their families. Census data are vital for delivering these council services effectively. For example, the Council’s Early Years, Children and Young People’s Services uses lower super output area
(LSOA) data from the 2011 Census together with the Income Deprivation Affecting Children Index (IDACI), a deprivation index that measures the local area’s proportion of children under age 16 that are living in low-income household, to prioritise initiatives and funding. More specifically, these services use: (i) LSOA census population estimates with IDACI to allocate around 85 per cent of all Children Centre funding in a targeted way; (ii) IDACI by LSOA for the allocation of deprivation supplements to providers of the free early education entitlement for three and four year olds; and (iii) LSOA census estimates to ensure sufficient childcare is available, assess take-up of Children Centre services, aid in developing and managing the childcare market, and plan for the implementation of the free entitlement for two year olds, all of which are statutory duties.

3.7 Conclusions

The census is a unique source of data that has a wide range of uses involving an extensive number of users. The 2011 Census was considered a success and the underlying reasons for that success were:

- the planning undertaken with the end-user in mind and being clear about the benefits that had to be realised – each Census Office shared a vision and an understanding about how the results were used and collectively worked towards fulfilling that vision;
- the programmes had agreed success criteria which were used to drive the design and decision making – continuous monitoring of progress towards goals, extensive testing and rehearsing of procedures and methods, responding to lessons learned and introducing new innovations to address lessons learned and changes in society and technology, as well as having sufficient resources and people with the right skills at the right time;
- recognition of the importance of stakeholder engagement and establishing partnerships with local authorities and community organisations and representatives; and
- ensuring that the confidentiality and security of the information collected in the census was a matter of the highest priority.

All of the Census Offices have undertaken a range of activities with users to promote and identify the benefits which flow from the census. These activities have been successful in broadening the use of the outputs and improving the Census Offices’ understanding of the uses of census data. For the first time, the Census Offices attempted to measure the scale of the benefits generated by the census and having an understanding of the wide range of users and uses of census data was critical to quantifying the benefits. Measuring the benefits of official statistics such as those collected by the census is difficult, as users of census information find it hard to put a value on their use. Therefore, it was important for the Census Offices to work closely with individual users and representative bodies (e.g. trade associations) to understand the use made and to explore what alternative data might be used.

ONS concluded an assessment of the 2011 Census benefits in England and Wales in January 2014. ONS identified and quantified the benefits for government departments, the wider public sector, local authorities, businesses and genealogy/family history. Many examples of uses of census data were gathered from other sectors but, because of the diverse and fragmented nature of some of these sectors, it was not possible to estimate total benefits for each sector. More detail on the methods used in estimating the benefits is detailed in Cope (2015). This assessment concluded that over a ten-year period, the costs for the 2011 Census were estimated at £482 million in cash prices. The re-valued 2011 Census benefits were estimated at £489.5 million.
each year, which equates to a payback period of around 14 months. The total benefits from the 2011 Census turned out to be significantly higher than the £720 million estimated in the 2007 business case. However, it was also noted that quantifying the benefits to central government in areas such as policy development and evaluation was particularly difficult.

Notes

1 In Scotland, the vast majority of census questionnaires were hand delivered by enumerators. Postal delivery was only used in around 5 per cent of cases covering remote areas.
2 The question on citizenship was not asked in Scotland.
3 The question on second residence was only asked in England and Wales.
4 The question on intention to stay was not asked in Scotland.
5 In Scotland, where hand delivery by enumerators was used in both 2001 and 2011, the size of field force reduced from around 8,000 to just under 7,000. Northern Ireland adopted postal delivery of census questionnaires for the vast majority of addresses (other than in rural Fermanagh), but maintained the model of an enumerator taking responsibility for a given enumeration district; its number of enumerators fell from around 2,500 in 2001 to around 1,500 in 2011.
6 Daytime populations in Scotland and Northern Ireland.

References


ONS (2012f) Proving the online census. 2001 Census Update, ONS, Titchfield.


