2011 Census

How the 2011 Census population estimates were obtained

December 2012

How the 2011 Census population estimates were obtained

Introduction

The main purpose of the census is to provide an accurate population count. Although every effort is made to ensure everyone is included in the census, inevitably some individuals are missed. This under-counting does not usually occur uniformly across all geographical areas or across other sub-groups (for example, by age and sex) of the population.

To fill the gap, NRS implemented a coverage assessment and adjustment process to estimate the population that was missed. In addition, this process identified and adjusted for the people who were counted more than once or who were counted in the wrong place. Carrying out this work allowed a census estimate of the entire population to be obtained. Detailed below are the various elements that were used to calculate the final 2011 Census population estimates included in this bulletin.

The methods were largely based on those developed by the Office for National Statistics (ONS). ONS have produced a full suite of methodology papers detailing the statistical theory and practical application of the methodology. They can be found here: http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-data/2011-first-release/first-release--quality-assurance-and-methodology-papers/index.html

Full details of the sizes of the adjustments which were made to the census results will be published as part of Release 1B.

Census Coverage Survey / Dual System Estimation

The primary source of 'missingness' was addressed using a Census Coverage Survey (CCS) in which about 40,000 households across Scotland were visited by trained survey interviewers. The CCS was a sample survey independent to the census, conducted six weeks after Census Day. Participation in it was voluntary.

Responses from both the Census and the CCS were matched and the known number of responses from each of these were used (in a statistical technique called Dual System Estimation) to determine the number of respondents missed for each of the sample areas in the CCS. The Dual System Estimators were then used to derive the population for a whole Processing Unit[1].

Communal Establishments

Small[2] communal establishments (CEs) were treated in the same way as individual respondents i.e. using a DSE to determine the missing people. This was carried out at Scotland level data (rather than PU by PU as the sample sizes would otherwise be too small) to produce accurate population figures to add in to the population estimate.

---

[1] A processing unit (PU) is made up of one or more neighbouring council areas (CAs) and consisted of approximately 500,000 respondents. CAs were grouped in to PUs for practicalities around data processing etc.
[2] Small communal establishments are those with less than 100 bed spaces. Large communal establishments are those with 100 or more bed spaces.

For larger communal establishments work was done to investigate the likely number of persons missed in the 2011 Census. This work included matching the number of actual questionnaires returned against the number issued by the CE manager, the number of available bed spaces and information gathered from administrative records. Where gaps were identified, the appropriate number of people were added in to the population estimate.

Between Household Bias Adjustment

The DSE method makes the assumption of independence between the CCS and Census so an individual's likelihood to respond to the CCS is not influenced by how they responded to the census. However, this assumption will not always hold due to the fact that i) households that are less willing and likely to respond to the census will also be less willing and likely to respond to the CCS; and ii) a household's likelihood of responding to one may be affected by whether or not it responded to the other.

To help overcome this bias, an alternative count of occupied households was calculated, based primarily on information gathered during the census field operation. This alternative count was fed in to the ratio estimation and was used (if necessary) as an inflation factor for the DSE that calculates people missed by both the census and the CCS.

Overcount

Particular groups of people may be included on one or more census questionnaires. An example would be students who completed a questionnaire at their term-time address but were also included in the one returned by their parents for the family home. Multiple responses for the same household were removed during the early stages of processing.

To produce an estimate of level of overcount, the census was matched to the CCS and to itself. This was used to produce conservative 'dampening' factors for broad population groups which modified the DSE and adjusted the census estimate.

Sample balance

The CCS sample would be expected to be an accurate representation of the overall population for the sampled areas. However, with every sampling process there is a risk that a sample may be an outlier amongst all possible samples. For example, the chosen CCS sample could have, by chance, drawn postcodes in an area where the census had managed to count everyone.

The CCS sample was assessed to determine if it was sufficiently representative when compared to all other possible samples that could have been drawn from the underlying population. This assessment showed that no adjustments were required to correct for sample balance.