

- Overview of methodology to derive the CORE datasets
  - All data are submitted through the on-line Continuous Recording system (CORE) by private registered providers and local authorities. CORE is a database where individual property lettings are recorded, alongside information about the property and the tenants. The data is entered onto 'logs' (one per letting) or as a bulk upload through a facility called eCORE. Entries are automatically validated as they are uploaded.
  - The contractor TNS are responsible for collecting data through CORE and validating the data.
  - The process of weighting and imputation of data is done by statisticians in DCLG. DCLG commissioned in 2013 ONS Methodology Advisory Service to advise on improving the methodology to derive the *Social Housing Lettings* statistics, with funding from the UK Statistics Authority Quality Improvement Fund. The main methodology changes were to estimate population totals for social housing lettings, by applying improved weighting methods, and address item non-response for the household characteristics (such as age, ethnicity and economic status) by imputing for missing values.
- Imputation
  - Some of the questions on the CORE questionnaire are not compulsory. In particular, data on household characteristics (age, sex, economic status, ethnicity and nationality) may not be available to the housing officer or may be refused by the tenant.
  - As recommended by the ONS Methodology Advisory review, DCLG uses the Canadian Census Editing and Imputation System (CanCEIS) software to impute data using a 'donor imputation' approach. This involves identifying variables that are related to the variables with missing data; and then choose the record with the closest data for the non-missing variables and copy the data for the missing variables.
  - Imputation is carried out to address item level non-response of key data on tenant characteristic, for both local authorities and private registered providers (with local authorities having a higher level of non-response). The imputation is done separately for general needs and for supported housing, to reflect the different demographic profiles of their tenants. The imputation is done in two steps, imputing first age, sex and economic status, and then nationality, on the missing variables for the household reference person.

- Income data is not imputed. Income has a higher non-response rate, particularly amongst the elderly. Therefore the data was deemed to be too incomplete and unreliable as imputed results may lead to biased estimates. Instead, to improve completeness of income data DCLG is working with TNS to improve the guidance and CORE form itself to encourage a higher response rate.
- Weighting
  - In order to obtain representative estimates across England and to be able to analyse data at local authority level, weighting is applied to adjust for record level non-response by local authorities. CORE has always captured all lettings by private registered providers in England, however local authorities have participated in CORE since 2004/5 on a voluntary basis. In the first year, only 24% of stock-holding local authorities participated, but the number of authorities participating has steadily increased since then with all authorities submitting some data for 2013/14.
  - Local authority weights are calculated by reference to the total number of lettings reported to the *Local Authority Housing Statistics (LAHS)* return, or its predecessor the *Housing Strategy Statistical Appendix (HSSA)*. Final LAHS data are not used as they are not published until December.
  - Weights are calculated for groups of similar local authorities, where these groups are defined by the ONS UK Output Area Classification ([link](#)). This National Statistics area classification is based on 41 detailed variables derived from the 2001 Census, many of which are directly relevant to the factors of interest for social housing (e.g. age, nationality, employment status).
  - For each geographical cluster, the weight is calculated for all responding local authorities as the ratio of the number of lettings reported to LAHS relative to those reported to CORE.
  - Neither imputation nor weighting are carried out for affordable rent lettings, mainly because it is still early in the Affordable Rents programme and so there are relatively few lettings. As the number grows the methodology may be extended to these lettings.
  - Ahead of the 2013-14 National Statistics release, DCLG changed the data used to derive the cluster weights to a more comparable dataset to the LAHS dataset. Instead of using the lettings totals based on the 'location' of the dwelling being let, data based on the local authority 'owning' that dwelling was chosen, because the LAHS data is also based on the ownership and not the location of the dwellings. This led

to lower weights and, in turn, to a smaller total number of dwellings at the national level. The weights for 2012-13 were also revised accordingly, with a new back series of datasets being released.

- For further information on the weighting and imputation methods, please see the report produced by ONS for DCLG on '*Improving Outputs on Social Housing Lettings*':  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/245601/Improving\\_outputs\\_on\\_social\\_housing\\_lettings.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/245601/Improving_outputs_on_social_housing_lettings.pdf)