

ONS Response to recommendations from the Independent Review of Coverage Assessment, Adjustment and Quality Assurance

1. Introduction

In October 2010 the National Statistician commissioned an independent review of the methodology for coverage assessment, adjustment and quality assurance of the 2011 Census population estimates. The review was published on 21 February 2011 with 23 recommendations.

ONS welcomes the report. The recommendations will help to improve the robustness of the methods and build user confidence in the census results.

The ONS response to each recommendation is in section 2 below.

In response to a number of the recommendations, ONS will publish a series of documents that provide more detail and clarity on some aspects of the methodology, building on the materials already published. A list of the additional documents that ONS will publish is in section 3

2. ONS response to each recommendation

Recommendation 4.1: Provide a more timely breakdown than was achieved in 2001, at the same level of detail as in 2001, of response rates for the CCS. This will enable users to assess the robustness of the DSE adjustment.

Agree. This information will be provided as supporting information when the census population estimates are published in July 2012. An outline of the material to be published with the census population estimates will be published in July 2011.

Recommendation 4.2: Consider capturing data about household non-contact and refusal rates by CCS interviewer and analysing the data to provide additional intelligence for use in QA and ratio estimation.

Agree. Information about non-contact and refusal rates will be collected for every household that does not respond to the CCS (Census Coverage Survey) through the completion of CCS 'dummy forms'. Further consideration is being given to capturing information from the CCS interviewer record books about

calling patterns and contact outcomes. However, this information would need to be keyed manually using additional, unplanned resource, so any exercise to capture this information would be prioritised in the areas where CCS response was lowest.

Recommendation 5.1: Set out in detail, ideally with some examples, the way in which DSE is applied to data from the Census and the CCS to produce estimates both for each age-sex group and also for each ethnic group, making it clear when the assumption of independence between the Census and CCS has been relaxed, explaining which levels of aggregation have been used and how different forms of post-stratification (for example, tenure and ethnic group) have been used to strengthen the results. The methods used for small CEs should be included.

Agree. We recognise the user need for this information to support their understanding and confidence of their census population estimates.

We will publish a large amount of detailed information, including the points specifically made in the recommendation, about the coverage assessment and adjustment and quality assurance process. This will be done in two tranches, one with the first release and one with more detail shortly after, to ensure that preparation for the second tranche of information does not delay the publication of the census population estimates.

ONS will publish the content and timetable for each tranche in July 2011.

Recommendation 5.2: Publish a detailed assessment of Census-CCS matching success rates for automatic and manual matching (and for each EA).

Agree. This will be published as part of the second tranche of supporting information described in the response to recommendation 5.1

Recommendation 5.3: Clarify whether the IHS is capable of being used for bias adjustment for persons in counted households.

Agree. Information from the IHS and other ONS Social Surveys will be used as part of the within household bias adjustment during the coverage adjustment methodology. An outline of the methodology for this component will be published in May 2011.

Recommendation 5.4: Given their use in subsequent outputs, give serious consideration to extending the quinary age groups to 90+, if necessary collapsing geography rather than age groups to secure sufficiently robust Census-CCS estimates amongst these higher age groups.

ONS will publish in May 2011 the age groups that will be used within the estimation process. The collapsing of geography and age groups during the application of the coverage adjustment methods depends on the numbers found within each of these strata during the estimation phase. As the numbers within a hard to count group, within an estimation area, within the CCS sample aged 90+ are likely to be very small it would take considerable geographic aggregation to reach acceptable levels of numbers to produce reliable estimates. Therefore ONS does not think it is prudent to commit to this now as this relies heavily on the size of the numbers found within the individual DSEs.

Recommendation 6.1: Clarify precisely how ONS will integrate the separate overcount propensities identified from its Census self-match and Census-CCS matching exercises as inputs to the DSE process.

Agree. This work is currently being finalised and will be published in May 2011.

Recommendation 6.2: If faced with timetable or resource pressures for estimating overcount, prioritise nationwide Census self-matching of those with a stated second address, or different address one year ago (not a currently listed ONS strata) above self-matching other strata less at risk of being a duplicate within the same GOR.

Agree. The timetable for processing and publishing the data relies on a number of assumptions about volumes and timings for each processing step. Should these assumptions be significantly wrong, thus incurring delays, then

this would be considered as a way of improving processing times. However if overcount is a significant problem, reducing processes which assess overcount may reduce the quality of the estimates.

Recommendation 6.3: Publish estimates of the components of overcount associated with each EA and, where possible, LA, including both removal of duplicate returns from within same postcode and the net overcount adjustment arising from Census-CCS matching.

Agree. This information will be provided as supporting information when the census population estimates are published in July 2012. An outline of the material to be published with the census population estimates will be published in July 2011.

Recommendation 8.1: Clarify the circumstances in which an LA fixed effect will be favoured, including the presence of atypical features within the EA such as substantial HE student residents.

Agree. An outline of the methodology for this component will be published in May 2011.

Recommendation 8.2: Publish the asymmetric confidence intervals derived from variance estimation in preference to symmetric confidence intervals, as those expert enough to use confidence intervals are likely to be expert enough to make use of this useful additional information.

ONS will publish asymmetric confidence intervals derived from the variance estimation process. The methods for estimating variance and deriving confidence intervals will be published in May 2011. This publication will also set out how confidence intervals will be reported, either alongside the census population estimates or as part of a methodological evaluation of coverage estimation and adjustment.

Recommendation 9.1: Finalise and publish details of the planned large CE imputation process in sufficient time to receive user feedback

before it is implemented and describe in the final documentation how imputation deals with small CEs.

Agree. The methods for imputing into small and large CEs will be included within the methods on coverage imputation to be published in May 2011.

Recommendation 9.2: Identify areas and population sub-groups with significant overcount and assess the extent of any bias that might be introduced by retention of duplicates, through a comparison of identified duplicate and missed records.

Agree. This analysis is part of the plans to evaluate the overcount component of the coverage assessment and adjustment methodology. This evaluation will be done after the census population estimates are produced and aims to evaluate how well the methods worked and identify improvements for future use of the methods. Any significant biases found during this evaluation will be reported to ensure that users have a clear understanding of the quality and any limitations of the census data.

Recommendation 9.3: As in 2001, publish imputation rates for each LA, age, sex, ethnicity and intention to stay category, so that expert users can take account of these when undertaking their own analyses of Census data.

Agree. An outline of the material to be published with the census population estimates will be published in July 2011.

Recommendation 10.1: In the light of our comments, identify those QA checks that are so strong that they are able to be used to improve the Census where necessary.

Agree. This will be included in the detailed QA methodology to be published in May 2011.

Recommendation 10.2: Provide a unified overview and detailed documentation of proposed QA methodology as soon as possible to allow users to understand its coherence.

Agree. ONS will publish a unified overview of the QA methodology in May 2011. Additional detail will then be provided in July 2011.

Recommendation 10.3: Undertake discussion in Spring 2011 to gain users' confidence in the QA and in the post-QA improvement procedures.

Agree. A range of engagement activities have already taken place over the past two years to build users' understanding and a further meeting is planned with the QA advisory group in late spring 2011 to discuss the outstanding elements of the QA methodology. In particular, among the key items remaining to discuss is the guidance from moving from core to supplementary QA. This work will be covered in one of the documents published in July 2011.

Recommendation 10.4: Confirm that knowledgeable, impartial support would be sought from the LA concerned during QA, when encountering a difficult and unusual pattern of discrepancies between the Census population estimates and detailed QA checks.

Agree. ONS recognises that where LA estimates are significantly different from other sources and contain unusual patterns that cannot be explained through the supplementary analysis, then support and input from the relevant LA would be valuable and will be sought. ONS would expect this to be the exception rather than the norm.

Recommendation 10.5: Use the strength of each QA check to prioritise QA work, giving low priority to work-intensive checks that are unlikely to be used in improving Census estimates. This priority should apply both in the current preparation for QA before summer 2011, and during the subsequent QA itself. In particular, the proposed procedures and datasets for post-QA improvement should each be fully specified and fully prepared prior to their use in QA. This does not preclude judgement and methodological refinements during the QA.

Agree. ONS has already acted upon this advice in refining the methodology for QA. Prioritisation on specific checks will ensure that resources are efficiently used allowing enough resource and time for a robust quality

assurance of all LAs but also to ensure sufficient resource for those LAs that require more detailed supplementary analysis.

The QA methodology paper to be published in May 2011 outlines the core checks for each LA. This list has been prioritised to reflect the recommendations of the review team.

The procedures and preparations for the post-QA improvement options are currently being developed and are on track for implementation when the QA begins in July 2011.

Recommendation 10.6: Prepare for and expect to undertake supplementary QA in LAs with the highest indications of concern from the HtC index and field reports.

Agree. The planning and preparation being undertaken for the supplementary QA is already focussed on the LAs that are expected to be the most challenging, and these were the focus of the QA studies conducted in 2010¹.

Guidance is currently being developed that will set out the conditions under which LA estimates would be expected to undergo supplementary QA. The guidance will take account of a number of factors including census field intelligence, which will also be used as part of the core QA of all LA estimates. A fuller paper outlining the guidance on moving from core to supplementary QA will be published in July 2011.

Recommendation 10.7: Ensure that the timetable for coverage assessment and QA is sufficient for the likely requirements of in-depth investigation for a minority of LAs as well as the national investigation that can only be finalised when all LAs have been processed.

Agree. The timetables for estimation and QA have been integrated (along with earlier processing stages) to be able to have sufficient time to assess the quality of the census results and improve them if necessary during the processing of the data. This timetable includes time for supplementary

¹ The final report of the QA Studies is available from: <http://www.ons.gov.uk/census/2011-census/process-info/data-quality-assurance>

analysis on a small proportion of LAs which require more detailed investigations.

Within the timetable there is an explicit period of time for considering the national estimates and any further improvements that may be identified as a result of the national level QA.

If additional time is required as QA reaches its final stages then ONS will consider delaying publication of the results.

Recommendation 11.1: The QA indicators for ethnic group should not be considered for adjusting the Census.

Agree. The QA sources will not be used directly as the basis for an adjustment. ONS is also considering a range of other related variables which can further support and enhance the ethnicity check, including country of birth and language spoken. Where these all point towards the possible need for an adjustment, a final view will be taken by considering an alternative stratification of the Dual System Estimate, using the ethnic group classification.

Recommendation 11.2: The plans to target large households and HMOs in early collection should be as comprehensive as possible. Where ethnically diverse LAs with substantial numbers of recent immigrants have not provided lists of HMOs for targeting by early collectors, these lists should be sought with urgency, targeted in those LAs for which published numbers of HMOs are high.

Agree. A final request for HMO lists was sent to LAs with published high numbers of HMOs. One of the tasks of early collectors, who start on 21 March, is to visit addresses which are known to be HMOs and areas where there is evidence of large households. The early collectors will offer householders continuation questionnaires, additional questionnaires for new households not on the Address Register and help in completing the questionnaire.

Every Census Coordinator Area will have at least one early collector, and there will be more early collectors in areas where there are large numbers of HMOs or large households..

3. Documentation to publish

The following table summarises the documentation that ONS will publish over the next few months in response to the recommendations of the report.

Recommendation number	Title	Description	Publishing date
5.3	Within household bias adjustment methodology	How ONS social surveys will be used to assess whether the assumption of independence within counted households in the DSE is correct, and if not how we will make adjustments to the population estimates.	May 2011
5.4	Age-sex groups to use in estimation	The default age-sex groups that will be used as estimation	May 2011
6.1 9.2	Assessing the level of overcount	How we will estimate overcount propensity and make adjustments to the estimates.	May 2011
8.1	Use of LA fixed effect estimator	How we will test whether we should use the 2001 approach for estimating LA populations rather than the simplified version	May 2011
8.2	Methods for estimating variance and deriving confidence intervals	How we will estimate variance and produce confidence intervals	May 2011
9.1	Methods for estimation and imputation within large and small CEs	What process we will use to estimate and adjust the output database populations of large and small CEs	May 2011
	Coverage Imputation Methodology	The methodology to adjust the output database	May 2011
4.1 5.1 5.2 6.3 9.3	Proposed content of first release of 2011 Census population estimates	Material to be published as the first release in July 2012	July 2011
10.1 10.2 10.3 10.5	Overview of Census Quality Assurance Methods	Document setting out quality assurance and checks	May 2011

10.6			
10.3	Guidance criteria for moving from core to supplementary quality assurance	Criteria for when supplementary quality assurance will be carried out	July 2011
