Options for making national adjustments to LA level

Owen Abbott, ONS
Dr James J Brown (UoS)

0 Summary

This paper outlines how to make adjustments at national level should the need arise to do so and 3 options for cascading down to local areas.

1 Introduction

We already have in place within the dual-system estimation process a number of approaches that estimate for and remove local sources of bias such as heterogeneity (or correlation) bias and overcount bias. The first involves information from the Census Household Frame (address register) to derive an odds ratio at the household level, and matching to other survey (and/or administrative data) to give an odds ratio for within household dependence. The second involves searching for duplicates, modelling the matched CCS data and making use of the LS analysis of duplication to derive weights that are applied to the DSE populations. These are not the only measures to make adjustments (see http://www.ons.gov.uk/census/2011-census/process-info/statistical-meth/2011-uk-census-coverage-assessment-and-adjustment-methodology---article-from-population-trends-137.pdf), and the strategy is to make these prior to assessing plausibility of the national population estimates.

At regional and national levels, there are additional data available to make an assessment of any underlying or residual bias that local adjustments were unable to remove. This paper outlines some of those sources, and the options for cascading any such regional or national adjustment to the local estimates. The paper does not outline or discuss the decision making processes for a national adjustment, as this is covered by the overarching Quality Assurance Methodology paper (see http://www.ons.gov.uk/census/2011-census/process-info/data-quality-assurance/2011-census---methodology-for-quality-assuring-the-census-population-estimates.pdf).

2. National and Regional Adjustments

2.1 Adjustment total

In order to make a national or regional adjustment, an adjustment total is required by age and sex. Denote these adjustments as A_a where a denotes the age-sex groups. This can come from a number of analyses for which work is underway, including analysis of the Longitudinal Study and through an agreement of the sex ratio that we would expect the census to provide. These are described briefly below. There may be other regional or national analyses that provide relevant adjustment totals for all or a selection of age-sex groups. However, it is important to point out that as yet none of these have been agreed as providing definitive totals or patterns.

2.1.1 Longitudinal Study

The longitudinal study will result in an analysis (the methodology for which is yet to be defined, reviewed and agreed) that will provide some information on the plausibility of the census estimates, by age and sex, at national level. In 2001, a decision was made to use this analysis to generate adjustments by age and sex. For 2011, until the methodology is fully outlined, it is difficult to state whether the same decision would be reached.

2.1.2 Sex ratios

If an agreed sex ratio is used to QA the national totals, the only option for making adjustments so that the census estimates are calibrated to the ratios is to adjust either the male or female estimates, keeping the other constant. To change both at the same time (by say moving up both males and females but at different rates) requires additional evidence from another source, which would need to come to a view as to what the total population should be. It is cleaner to fix the sex ratio first, and subsequently uprate the total population to the other external evidence (i.e. the sex imbalance is corrected first by (for instance) adding in males, then a further adjustment is made to the total population that keeps the calibrated sex ratio constant).

Where this has been applied in other countries, the default strategy has been to fix the female total and adjust the male estimates. This is because in most cases the sex ratio in the census estimates is too low and this is fixed by adjusting the male estimates upwards so that the resulting sex ratio is in line with the agreed evidence.

2.2 Options for cascading adjustments

Once a national or regional adjustment has been identified and agreed (for all or a selection of age-sex groups), there is then a need for a method to cascade this adjustment down to Local Authorities. This then allows the revised estimates to be fed into the relevant processing stage (Coverage estimation) and thus the Census outputs to reflect these adjustments. There are essentially 3 options, described below. Each of these options as described can make adjustments to all LAs, or to a subset of LAs, depending on what is required. The decision as to whether all or a subset of LAs would need adjustment is outside the scope of this paper but will inevitable rely on judgements as to whether certain LAs estimates are plausible without further adjustment.

Option 1 – proportional to population size

A simple apportionment is made on the basis of population size. Thus if our original estimates are $\hat{T}_{al}$ where l denotes the local authority then the adjusted estimates $\hat{T}^{*}_{al}$ are then:

$$\hat{T}^{*}_{al} = \hat{T}_{al} + A_a \times \frac{\hat{T}_{al}}{\sum_{l=1}^{n} \hat{T}_{al}}$$

This is equivalent to applying a proportional uplift to each age-sex group. This is not necessarily a good method to use, as it will put in more population in absolute terms where lots already exist.

Option 2 – proportional to coverage adjustment

The second option is to apportion the adjustment based on the size of adjustments already made through the coverage processes. Thus if the census counts are X_{al} then the adjusted estimates $\hat{T}^{*}_{al}$ are then:

$$\hat{T}^{*}_{al} = \hat{T}_{al} + A_a \times \left(\frac{\left(\hat{T}_{al} - X_{al} \right)}{\sum_{l=1}^{n} \left(\hat{T}_{al} - X_{al} \right)} \right)$$

This then applies a proportional uplift to each age-sex group according to the adjustment already made. Thus it will put in more population in absolute terms where lots of adjustment has already been made. Hence areas with lower coverage will have more adjustment. This could be finessed by doing this by the HtC index as well.

Option 3 – proportional to missed by both

The third option is to apportion the adjustment based on the estimated size of the missed by both cell arising through the coverage process. This is essentially then adjusting the odds ratios used to drive the DSE bias adjustments. The rationale for this method is then clearer – it implies that there is some residual within household bias that has not been adjusted for, and thus the DSEs are lower than they should have been. This makes the assumption that there is no remaining between household bias (if there is, then the household based DSE bias adjustment should be modified first to recalculate that adjustment before applying this (revised) adjustment). This is the methodology proposed (but ultimately not used) in 2001 for making adjustments based upon a sex ratio.

One issue with this adjustment method is that it has to be first cascaded down to estimation areas (represented by g), as that is the level at which we can obtain robust estimates of the missed in both cell.

Thus if $\hat{n}_{00ag}$ are the estimated 'missed by both' cells from the DSEs, after the various DSE bias adjustments have been applied (so the formulae for this is not derived here) then the adjusted estimates $\hat{T}^{*}_{ag}$ are then:

$$\hat{T}^{*}_{ag} = \hat{T}_{ag} + A_a \left(\frac{\frac{\sum_{sample} n_{00ag}}{\sum_{sample} X_{ag}} X_{ag}}{\sum_{g} \left\{ \frac{\sum_{sample} n_{00ag}}{\sum_{sample} X_{ag}} X_{ag} \right\}} \right)$$

The LA level adjustments are then derived from the regular LA estimation methods used in the coverage assessment methodology. Thus, the Estimation Area estimates are adjusted and fed into the estimation system. The LA estimates can be approximated outside of the system (to check for plausibility before committing to the adjustment) by apportioning out the EA level adjustment according to the LA patterns seen in the unadjusted results.

This adjustment will put in more population in absolute terms where lots of adjustment has already been made, but this will be more in line with where both census and CCS coverage was poor, so it take account of some of this uncertainty in the estimates. This could be finessed by doing this by the HtC index as well.

2.3 Discussion

Of the three options presented above option 3 is probably the one that fits within our framework the best, even though it is a more complex calculation and may be difficult to understand by users. However, it may provide a better set of adjustments than the other options. Further work to show how these three options differ through a numerical example may help to make a decision as to which is the default method.