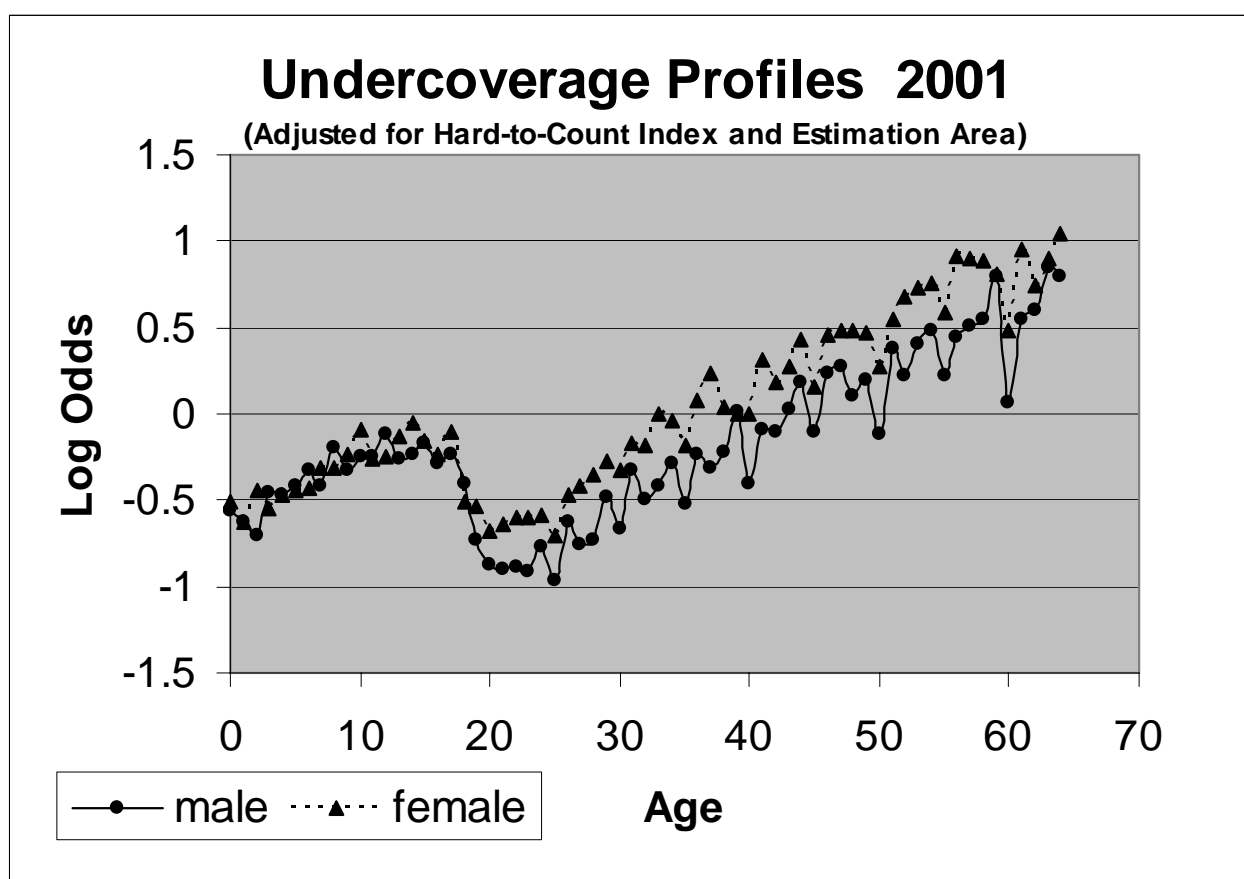# Age Sex Groups for use in Estimation

Alan Taylor, Nargis Rahman and James Brown

The estimation methodology allows for the age-sex groups used in the estimation process to differ from those in the outputs. In 2001 36 five year age-sex groups were used plus an estimate for children under 1. These were closely linked to the output categories of the mid-year population estimates.

The primary characteristic required of the grouping used for estimation is that the groups should be homogeneous in terms of their Census coverage probabilities as estimated from the CCS. This ensures appropriate estimates for the constituent age groups are calculated. The validity of this assumption for the chosen default age-sex groups will be monitored during the processing of 2011 Census data. A default categorisation based on the overall coverage profile across the whole country as seen in 2001 will be used initially. Criteria for aggregating single years of age into age groups were based on analysis of rates estimated using a logistic regression model. This model was estimated conditional of the effect of the Hard-to-Count area classification and the estimation area geography used in the 2001 Census (see Figure 1 for results to Age < 65).

Figure 1: Under Coverage Profiles for 2001 for Males and Female (age range 0-64)

The default grouping will be amended and estimation and adjustment rerun if the assumption of homogeneity of the default grouping needs to be revised based on observed patterns of coverage in the 2011 Census. Age-sex groups identified from this analysis will differ from the 2001 categories in the under 20 age group. The default age-sex grouping will be: **Male 0-2; 3-7; 8-17; and equivalent female groups; Male and Female combined 18; Male 19-24; 25-29; 30-34; 35-39**

1

then five year age groups to 85-89 and then 90+ and Female equivalent groups to male for these age ranges. This approach will help reduce heterogeneity bias in the DSE provided that the sample sizes are sufficient within these groups. Based on experience from 2001, the expectation is that the sample sizes in the 90+ groups are not likely to be large and will be collapsed to create 85+ or even 80+ groups. However, as noted in the ONS response to the independent review, we will consider whether a geographic collapsing would provide better estimates for the 90+ grouping. This would involve analysing whether the assumption made when collapsing age-sex groups causes any detectable bias through examining geographic collapsing for these age-sex groups.