Social Policy in a Cold Climate

Research Note Series

CASE
CENTRE FOR ANALYSIS OF SOCIAL EXCLUSION

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

RN003 – April 2013

Post-censal household estimates for small areas

Alex Fenton

Background

This is one of a series of short papers which explain conceptual or methodological approaches underpinning analysis undertaken in CASE's research programme *Social Policy in a Cold Climate* (SPCC). SPCC is designed to examine the effects of the major economic and political changes in the UK since 2007, particularly their impact on the distribution of wealth, poverty, inequality and social mobility. It also examines geographical variations in policy, spending, outputs and outcomes, with a particular focus on London. The analysis includes policies and spending decisions from the last period of the Labour government (2007-2010), including the beginning of the financial crisis, as well as those made by the Coalition government since May 2010. The programme will conclude in 2015, with publication of a final volume. Interim reports will be published in 2013/14, and made available online at http://sticerd.lse.ac.uk/case.

Social Policy in a Cold Climate is funded by the Joseph Rowntree Foundation and the Nuffield Foundation, with London-specific analysis funded by the Trust for London. The views expressed are those of the authors and not necessarily those of the funders.

Introduction

The Censuses of Population in the constituent countries of the UK provide counts of households at small area level and above every 10 years. All four countries in the UK also produce annual household projections and estimates, although by different methods (National Records of Scotland 2011). However, none produces such estimates for geographic units smaller than council or local authority areas. This note discusses methods for producing annual household estimates for "neighbourhood"-scale areas (LSOAs/Datazones) in England, Wales and Scotland that are consistent with the official local-authority level household estimates.

JRF JOSEPH ROWNTREE FOUNDATION

Nuffield Foundation

Trust for London
Tackling poverty and inequality

Our requirement for estimates of household numbers in small areas in years where Census data are not available arises from work on estimating poverty rates for small areas in order to understand the changing spatial distribution of poverty in UK cities. A family of measures has been developed, called UMBR: the *unadjusted means-tested benefits rate (*Fenton 2013). UMBR is intended to be a small-area proxy for income poverty as it is conventionally understood and measured by national surveys. In principle, poverty is commonly understood as a situation that affects families or households rather than individuals, and so the household or family is a common unit for the measurement of poverty rates. Practically, many poverty proxies, such as means-tested benefits, only identify one member of a household or family.

It is acknowledged that the household is not a perfect denominator, since its numerator (means-tested benefit claimants) includes some people who are not in households. Residents in nursing homes and other such institutions, for example, may claim some welfare benefits. However, it is preferred to the alternatives due to its consistency with national standards and for empirical reasons. UMBR denominated by counts of *adults* produced greater error and bias when compared to econometrically modelled neighbourhood poverty rates than UMBR denominated by *dwellings*, taken to be a proxy for household numbers. This is partly because there are systematic variations between places in the average number of adults in each household or benefit unit. The public small-area data on benefit claimants does not enable the adult partners of benefit claimants to be identified.

However, counts of dwellings (taken from Council Tax Band records) are not a perfect proxy for households. They create bias because there are variations in the relationship between dwelling counts and household counts. In particular, a dwelling count denominator will tend to understate poverty rates where there are many vacant dwellings in a neighbourhood. This is of concern for areas of high poverty and low housing demand (for example, parts of Teesside, and smaller ex-mining or ex-manufacturing towns), and for some coastal areas where many dwellings are second homes. Comparisons of the 2001 Council Tax Band statistics and the 2001 Census results showed that there were also some significant discrepancies between the Census's count of 'household spaces' and the Council Tax records of 'dwellings'. This is partly due to small differences in definitions, but the largest discrepancies are due to known failures in the enumeration of a small number of districts in the 2001 Census.

Therefore, methods for estimating the numbers of households in each neighbourhood, on an annual basis between Censuses, were investigated. This work was conducted in 2012, prior to the release of 2011 Census household counts. At the end of the note, we compare our 2011 estimates by our preferred method, with the 2011 Census counts, which have subsequently been released.

Overview of estimation methods

All four countries in the UK produce inter-censal estimates or projections of the number of households, although using different methods, and to the level of local authorities only. ONS (for England and Wales) and GROS (for Scotland) also produce mid-year estimates of the total population in small areas, by age band and sex, although, again, by different methods. Some official publications discuss methods for estimating LSOA household populations; the Environment Agency, for example, recommends dividing household population by estimated average household size (The Environment Agency 2012, chap.9). However, there are no published estimates of households or household population for small areas.

Two methods were investigated. The dwelling-change method starts from the number of households counted in the 2001 Census, and rolls these forward, making adjustments using administrative data on changes in the dwelling stock in each small area. The demographic-headship method starts with the small-area population estimates, and uses the proportion of each age-sex band believed to be a household head to derive a total number of households. In both cases, the results are constrained so that they sum to the official household projections for higher geographic units, such as local authorities and regions.

For completeness, both methods that were investigated are described below. The demographic-headship model is the one that was eventually used to produce the household counts provided with the UMBR dataset. It was preferred for several reasons: it produces fewer erroneous results; it is consistent with techniques used in official national and sub-national household projections; and, by using ONS/GROS small-area population estimates, it draws indirectly on a wider range of administrative data.

Dwelling-change method

This method is what is referred to as an ‘additive-change’ method; it consists of:

1. Starting with the Census counts of occupied and total household spaces.
2. Adjusting these to compensate for under-enumeration in the 2001 Census so that the total households for each district is consistent with the household estimates
3. Rolling these numbers forwards for 2002 to 2011, using year-on-year changes in Council Tax Band statistics to approximate the effects of construction and demolition, and then constraining each year's figures to the official LA-level household projection.

Base statistics

The Census is considered to provide counts of households and dwelling spaces that are consistent with the purpose at hand. Counts of total, vacant and occupied household spaces for LSOAs come from Census Table KS16. The base count of households is the total number of occupied household spaces. A 'household space' in the 2001 Census is not completely equivalent to a 'dwelling'; a dwelling may contain multiple household spaces which share kitchens or bathrooms (bedsits, for example). However, the number of shared household spaces is relatively small (c 80,000 in England and Wales, relative to a total of 25m), so this distinction is not maintained; household spaces are referred to as 'dwellings' here onwards.

The 2001 Census failed to enumerate some local authorities accurately, producing serious undercounts of households and dwellings in Westminster and Manchester, and some under-counting in other urban areas, particularly in London (Office for National Statistics 2004). This is dealt with by adjusting the count of households and dwellings upwards for all local authorities so that the household total for each LA is consistent with the 2001 household estimates, which are based on the mid-year population estimates corrected for the Census under-enumeration.

The adjustment is applied across all LSOAs in an LA. In districts where the sum total of households in LSOAs according to the Census is **less** than the relevant household estimate for 2001, an upward adjustment is applied to LSOAs where the count of all household spaces is less than the Council Tax (CT) records of number of dwellings. The "missing" households are distributed to each LSOA based on its share of all such missing dwellings within that LA. No households are added to LSOAs where the number of household spaces in the Census tables is the same as or greater than in the CT records. The occupancy rate (the number of households divided by the number of dwellings) is held constant. In districts where the sum total of households in LSOAs in the Census is **greater** than the household estimate for 2001, the number of households in each LSOA is reduced proportional to the share of all households in that LA that were living in that LSOA.

In most cases, the adjustment is nil or very small. Downward adjustments are no greater than 1%. There are some large upwards adjustments, especially in Westminster and Manchester. 191 LSOAs (about 0.5% of all LSOAs in England and Wales) have their household populations increased by more than 10%. This is consistent with the report on Census under-enumeration which found specific neighbourhoods to have been poorly enumerated (for example, those which had undergone development).

Neighbourhood-level change 2001 – 2011

The CT Band dwelling statistics are not consistent with the Census-defined household spaces. However, it is assumed that *changes* in CT Band dwelling counts approximate changes in

household spaces as counted by the Census. For each LSOA, for each year, the CT Band dwelling count is compared to the previous year to get three measures:

1. **Gross additions**: the sum of all positive changes within each band
2. **Gross subtractions**: the sum of all negative changes within each band
3. **Net Change**: the net change in total number of dwellings (= additions - subtractions)

As a simplified example (there are in fact eight Council Tax Bands, or nine in Wales), consider the figures for an LSOA in 2001 and 2002:

	Band A	Band B	Band C	Band D	Band E	Total
2001	50	100	100	50	20	320
2002	40	100	120	55	20	335
Change	-10	0	+20	+5	0	+15

The gross additions are +25 (+20 +5), the gross subtractions are -10, and the net change is +15. Changes happen for a variety of reasons: appeals against banding; the demolition and construction of dwellings; conversions to and from non-residential use; conversions of multiple dwellings to single dwellings and vice-versa. Data published by the VOA in 2011 suggest that the majority of changes reflect real changes in dwellings, rather than administrative re-bandings.

Rolling forward, applying changes, and constraining

Starting with the adjusted Census count of households and dwellings, these changes are applied each year, as follows:

1. The household count is increased by 1 for every gross addition to the dwelling stock
2. The household count is decreased by (1 * the occupancy rate) for every gross subtraction to the dwelling stock. In other words, an area's vacant dwelling rate is maintained – it is assumed that some demolished buildings are unoccupied.
3. The dwelling count is increased or decreased by the net change.

This gives a new provisional household and dwelling count for each LSOA. The household count for each LSOA is then constrained so that the total for LSOAs in each LA is consistent with that year's official household estimate. The adjustment is made proportional to the number of vacant dwellings in each LSOA. This means that when the LA-level household estimates are higher than the unconstrained total (i.e. there are more additional households than neighbourhood-level changes in dwellings suggested), neighbourhoods with more vacants tend to fill-up faster than those with few, and no neighbourhood can ever have more households than dwellings. When the

opposite is true, areas with higher vacancy rates will empty-out faster. Either way, the broad relative pattern of dwelling vacancy is maintained. Demolition and rebuilding will tend to reduce the vacancy rate of a neighbourhood relative to its local area, as possibly unoccupied dwellings are replaced by new dwellings that are assumed to be occupied.

Demographic-headship method

Headship rates (or household-representative rates) are used as the first stage in the English and Scottish official household projections (see, for example, DCLG 2010, p.5ff). Censuses and surveys identify one person within a household as the 'household head' or 'household reference person' (HRP). The term and its definition have varied over time. So long as a consistent definition is used within a single projection, the precise rules for identifying who in a household is 'head' or 'HRP' does not matter greatly, since the identity of interest is the one-to-one correspondence between the number of households and the number of HRPs. There are systematic differences by age, sex and marital status in the propensity of individuals to be HRPs: for example, older men are most likely to be HRP of their households. From this comes the idea of a headship-rate, or HRP-rate: the proportion of an age-sex group who are HRPs.

In overview, the method thus involves:

1. Starting with the mid-year population estimates for small areas produced by ONS and GROS
2. Deducting from each broad age-sex group the number of people thought to be living in communal establishments, and thus not part of the household population.
3. Multiplying the number of people in each age-sex group by the proportion of that group thought to be the reference person for their household (HRPs). The proportion of each band that are HRPs is derived from small-area Census data, and from the Labour Force Survey.
4. Summing for each small-area the numbers of households in each age-sex band.

It is considered an estimation, rather than a projection method, since it primarily uses available empirical data rather than projecting trends from past data.

Base Statistics

Two census tables are used to provide details on each small area. CAS001 provides counts of people by age and sex and whether in a household or communal establishment. CAS003 provides counts of HRPs by age and sex. From these tables two sets of statistics are extracted:

1. The local (LSOA/Datazone) HRP-rate, by age and sex band; this is the number of HRPs in age/sex band, divided by the number of people in households in that age/sex band.

2. The local count of people in communal establishments, by age/sex band

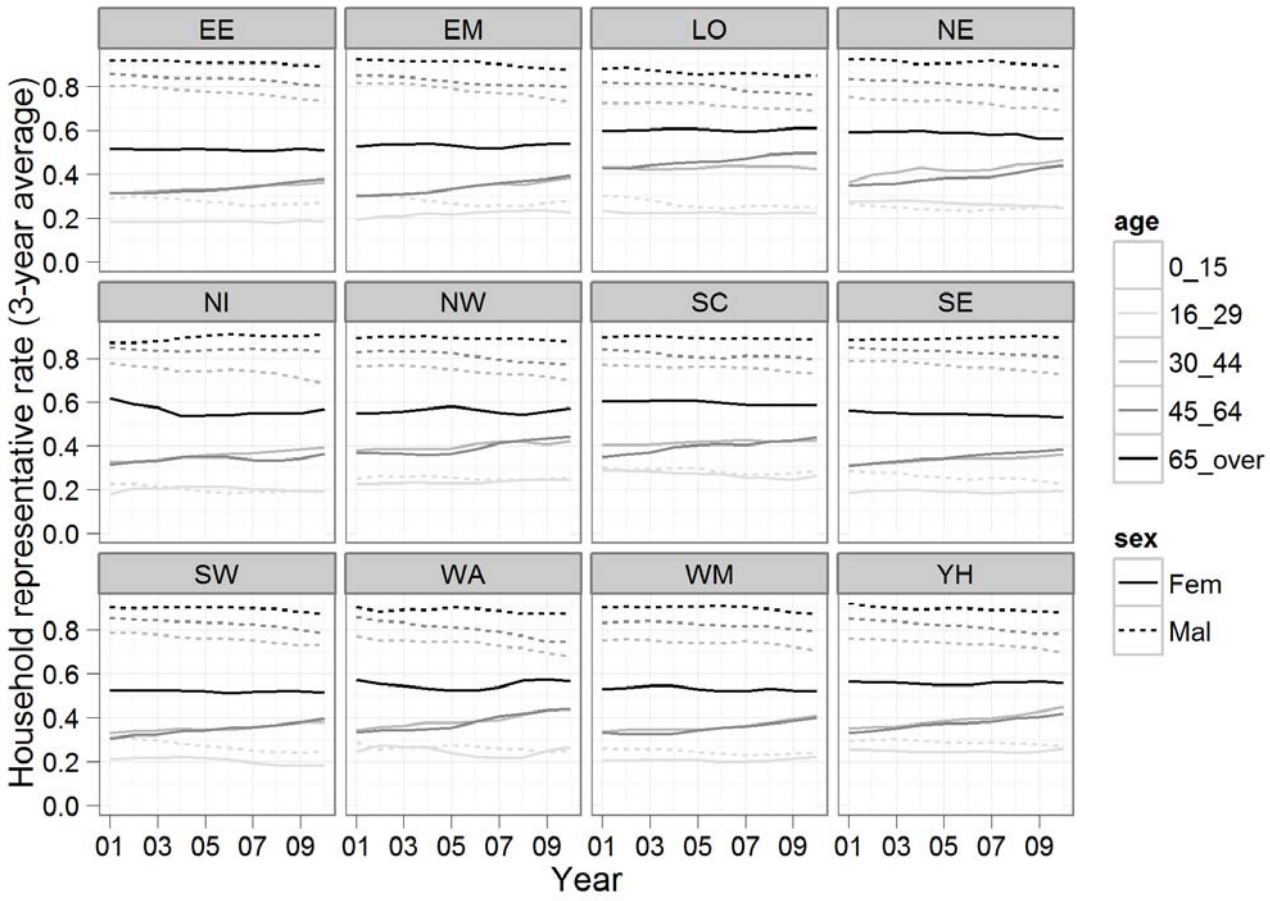
The small-area population estimates from ONS and GROS are collated into a single file covering all the years of interest (2001 to 2010) and with consistent broad age bands (0 to 15, 16 to 29, 30 to 44, 45 to 59/64, 60/65 and over). The small-area population estimates are already calibrated to local-authority-level mid-year population estimates. Correspondence with the Population Estimates team at ONS confirmed that this corrects for under-enumeration in the 2001 Census, and specific corrections were applied to small areas (specific neighbourhoods in Manchester, for example) that were believed to have been poorly enumerated.

Headship Rates

Since the household estimation period runs from 2001 to 2011, we have to consider differences in headship rates over both space and time. It is expected that there is variation over space in HRP-rates for a given age/sex band. This is partly because the housing in different places is likely to cater to different types of household, with different heads, and partly because there are differences between places in the typical marital status of adults in a given age group. Marital status is directly included in the national projections, but is implicit in the LSOA model.

The Labour Force Survey (Office for National Statistics, Social Survey Division 2010) shows that over the 2000s there were also considerable changes in the headship rates for age/sex bands, and differences between regions. Broadly, they show falling headship rates for younger and middle-aged men, and rising headship rates for women. These reflect various processes, such as the worsening affordability of open-market housing and increasing numbers of divorced, widowed or never-married women. The following chart illustrates the trends for different groups in different UK regions, as observed in the LFS.

Figure 1 : Household headship rates by region, sex and broad age group, 2001-2010 (Labour Force Survey). Male headship rates are shown by dotted lines, female by continuous lines.



The further the estimation date is from the Census date, the less confident one can be that the local variation persists, both because areas change characteristics (albeit, usually rather slowly) and because of the broad changes. Therefore two initial sets of household counts are made for each LSOA/DZ for each year, one applying local age/sex HRP-rate from the Census to that year's population estimate, and one applying the regional three-year moving-average HRP-rate for that age/sex group. The final household estimate for each area is a weighted average of the two. The closer the estimation date is to the year of the Census, the greater the weight given to the Census-based estimate. In 2001, the weight is 0.9/0.1 in favour of the Census; in 2010, the weight is 0.4/0.6 in favour of the LFS rates. The difference between the two estimates is greatest in city-centre LSOAs which have unusually high rates of headship for young adults in 2001.

These averages are then constrained to the official household projections for their local authority for that year, by simply keeping a constant share of the local authority's households in each LSOA/DZ and scaling up or down as necessary. The scaling is small in most districts in 2001 (maximum +/- 3%); by 2010 the greatest scaling is +/- 10%. Since there are no small-area population estimates for 2011, 2011 household estimates were derived by constraining 2010 LSOA shares to the 2011 official household projections.

Demographic Headship Model Preferred

A full comparison of the results produced by the two methods is beyond the scope of this note, and of limited value. The dwelling-change method produced a number of erratic results, including nil counts and some counts substantially below the minima suggested by other administrative records. The demographic-headship method benefits from ONS and GROS's methodological work and multiple sources that go into the small-area population estimates. Its technique is consistent with that taken in official household estimates for higher geographic levels. It also produces fewer erratic results. For these reasons, this is our preferred method of estimation for the UMBR dataset.

2011 and beyond

Since the work on which this note is based was completed, small area household estimates have been made available from the 2011 Census. Comparing the demographic headship model estimates with the Census estimates gives an indication of their accuracy, although the Census itself is of course prone to under-enumeration.

The table overleaf shows the difference between the household estimates for LSOAs in eighteen major English and Welsh cities (including their hinterlands)[1] and the Census counts. For this comparison, we exclude 519 LSOAs (2.5%) across the eighteen cities considered, because they did not exist on the same boundaries in both 2001 and 2011. These LSOAs are likely to have been ones which experienced the greatest population change.

[1] These cities are included because they are the subject of an analysis in which we are using the household estimates as a basis for estimating changes in the spatial distribution of poverty at the small area level within British cities. The LSOAs within them make up about two-thirds of LSOAs in England and Wales. For definition of the cities, see research note RN004 in this series. For details of the small area poverty measure (UMBR) see SPCC Working Paper 01: Small Area Measures of Income Poverty (Fenton 2013)

Differences between Modelled Household Estimatesfor 2011 and Census Estimates

City	Maximum Underestimate	Maximum Overestimate	25th percentile	50th percentile	75th percentile	N (lsoas)
Birmingham	-66%	40%	-6%	0%	6%	1774
Bristol	-61%	26%	0%	5%	10%	554
Cardiff	-40%	39%	-4%	2%	7%	451
Derby	-43%	24%	-2%	4%	8%	242
Hull	-39%	29%	0%	5%	9%	305
Leicester	-37%	37%	-5%	1%	7%	502
Liverpool	-138%	32%	-8%	0%	6%	976
London	-98%	46%	-5%	2%	8%	8552
Manchester	-64%	38%	-7%	0%	6%	1706
Nottingham	-25%	58%	-2%	3%	8%	478
Plymouth	-41%	27%	-2%	4%	8%	230
Portsmouth	-55%	27%	-7%	0%	5%	376
Sheffield	-27%	29%	-3%	3%	7%	508
Southampton	-73%	39%	-4%	2%	6%	409
Stoke	-39%	25%	-7%	0%	5%	335
Swansea	-35%	19%	-4%	2%	7%	290
Tyneside	-65%	35%	-5%	1%	6%	934
West Yorkshire	-52%	66%	-3%	2%	8%	1342
All	-138%	66%	-5%	2%	7%	19964

The table shows that the median difference between the counts is typically slightly higher than the Census estimate, between 0% and 3%. In four cities, Hull, Derby, Bristol and Plymouth, the median difference is higher, 4-5%. Overall half of the differences lay between five per cent lower than the Census estimate and seven percent higher. When we use the Census estimates to calculate our small area poverty rate (UMBR), rather than the modelled estimates, the difference is typically not more than one percentage point in either direction. The correspondence between the estimates and the Census is sufficiently close, in our view, not to re-base our calculations of poverty rates for the years between 2001 and 2011 on a retrospective, Census-based, household estimate. However, the table also shows that there are outlying LSOAs for which there was a substantial difference between the modelled estimates and the Census. One LSOA had, according to the modelled estimates, 138% fewer households than counted at the Census, and another 66% more. These cases could be due to inadequate correction for under-enumeration at either Census, or to dramatic population changes (due to demolition or new home building, which could also change headship rates) that were not picked up in inter-censal small area population estimates. When conducting specific city analyses, care needs to be taken over the interpretation of change in these outlying areas.

Households or a similar denominator for UMBR will need to be produced from 2011 to 2014 for the Social Policy in a Cold Climate research programme. It will be possible to apply the headship method described to base data from the 2011 Census; the relevant tables will be available. A potential development would be to estimate the numbers of 'families' rather than households, since 'families' as defined in the LFS and Census are closer to the definition of benefit-units.

The sources will benefit in several ways from the results of the 2011 Census. Official household projections for higher areas are to be updated to a 2011 base. Also, for the first time, small-area counts of population and households will be produced to the same geographic units as in a previous Census. This will permit new evaluations of methods for producing inter-censal estimates; ONS has underway a project for this which will report in 2013, and which will set out the basis for small-area population estimates from 2011 onwards.

References

DCLG (2010). *Updating the Department for Communities and Local Government's household projections to a 2008 base: Methodology*, London: Department for Communities and Local Government. Available at: http://www.communities.gov.uk/documents/statistics/pdf/1780350

Fenton, A (2013) *Small Area Measures of Income Poverty*. SPCC Working Paper 001. London: Centre for Analysis of Social Exclusion

National Records of Scotland (2011) *Household projections across the United Kingdom: Technical Report*, Cardiff: Statistics Directorate, Welsh Government. Available at: http://wales.gov.uk/docs/statistics/2011/110830hsetechreporten.pdf

Office for National Statistics (2004) *2001 Census: Manchester and Westminster Matching Studies Full Report*, London: Office for National Statistics. Available at: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/local-authority-population-studies/2001-census---manchester-and-westminster-matching-studies-full-report.pdf [Accessed September 28, 2012].

Office for National Statistics, Social Survey Division (2010). *Quarterly Labour Force Survey Household Dataset, April-June 2010* 2nd ed., Colchester, Essex: UK Data Archive. Available at: http://dx.doi.org/10.5255/UKDA-SN-6575-1

The Environment Agency (2012). *Methods of Estimating Population and Household Projections: Update 2012*, Bristol: The Environment Agency. Available at: http://publications.environment-agency.gov.uk/PDF/GEHO0612BWSA-E-E.pdf