

Whither / wither the census?

Paul Norman

Beyond 2011

Those of us who carry out population geography, applied demography or social science research have probably taken for granted that there will be a next census and we will have access to small area data. The only wrangling most people are aware of are about aspects like, “Will there be an income question?”, “How can we collect data from all those awkward people who steadfastly manage to avoid being enumerated?” and “How will confidentiality be preserved?” Actually, the taking of a census and the subsequent availability of data has not been a ‘given’ in the past and is especially not so with regard to the prospect for a 2021 Census.

In brief, we have a Census Act (1920) which provides the legislative framework for the taking of a census but each census requires specific secondary legislation. The first stage is a White Paper which contains the proposals for the census content and the data collection / dissemination strategies. Information in the White Paper is a result of user consultations by the Office for National Statistics (ONS) and the content is presented by the National Statistician (Jil Matheson) to the Minister for the Cabinet Office (Francis Maude) and the process continues via the Government Home Affairs Committee and then to Parliament which ultimately decides whether we have a census and its content.

Within this process, there is a need for ONS to make a business case for a 2021 Census or for alternatives. Essentially, this is a cost-benefit statement. Does it cost less to collect and disseminate the data than the value of the benefit of having those data? The ONS Beyond 2011 project is investigating alternative census strategies to the traditional model and whether alternative administrative data sources are viable. The assessments have two broad dimensions: counting the population by age and sex and characterising the population attributes (e.g. economic activity, ethnic group, tenure, household living arrangements, general health, etc.). There are equivalent enquiries in Scotland and Northern Ireland. The case for small area data will be the hardest to make.

To date, on a decennial basis we are used to having tables of cross-tabulated census data available from national down to local level. On an annual basis, we have access to administrative data for local geographies (useful sources but only univariate) and to large scale survey microdata such as the Labour Force Survey (rich in attribute information but poor in geographical detail).

For the future, there are a variety of choices being looked into but two versions can give us what we are used to, a traditional census and linked administrative records. Both can provide cross-tabulations at small area level. A census is what we know about and instinctively what most of us would opt for. Having linked records is an amazing opportunity since we could have more data richness than a census and have more frequent availability. There is, however, an expectation in government that we will **not** have a 2021 Census since it is too costly and we can expect a huge fight to get political and public acceptance of linked administrative data (understandably) if this was the advised option. This means that come 2021 we may have neither a census nor linked administrative data, in other words, no small area data.

There need to be some very compelling case studies demonstrating how we (and that includes all of you to whom this applies) have used small area counts and attributes and that if you hadn't had access to data collected in some way (i.e. not modelled) that something would have been impossible to do or so substandard that the utility of the work would have been compromised. Is there something which your organisation is legally required to do? Is there a service you provide which would fall apart without small area data? This isn't about something which we find interesting which we have done before and want to repeat, this is about things at a local level which society **needs** us to do.

Census as infrastructure

I want now to highlight an aspect which does not seem to form part of the Beyond 2011 agenda but something which we might take for granted. If we lose the traditional decennial census, then we no longer have the frameworks the census provides which is much more than just a 10 yearly update of data.

Taking a census provides a deadline. Planning for the census has timing targets built in. We have consultations on topics, questions and outputs, there are rehearsal and test censuses and census day itself. The census is also the point that we take stock, look back and re-base

previous mid-year population estimates in part because we believe that the census data are of better quality than an estimated population. This is not to say that an alternative statistical system cannot have a ten yearly re-grouping if deemed useful, but a time plan needs to be devised otherwise we might just drift.

Taking a census allows geographies to be defined. In the main, larger census geographies have been aligned with other administrative processes (regions, counties, local government, health administration, etc.) for the boundary systems existing at the time of the census. At sub-local government level, electoral wards have traditionally been used for analyses along with smaller Enumeration Districts. For the 2001 Census, statistical geographies were devised aiming to provide both geographic and socio-demographic detail including Output Areas (OAs) and aggregations of these to Lower and Middle Super Output Areas (LSOAs & MSOAs). The criteria that define these geographies include counts of people and households, devised so that there is a minimum that is large enough to preserve respondent confidentiality and a maximum which keeps the areas as small as possible. One of the knock-on advantages of these OA / SOA geographies is that they have remained frozen during the 2000s with numerous administrative datasets being made available (for LSOAs in particular) so that time trends (e.g. of unemployment) can be tracked. A similar approach has been taken for the dissemination of 2011 Census data with only minimal boundary changes from the 2001 definitions. The analysis of change over time, a major use of census data, is enabled by these geographical definitions.

If we do not collect census data at an individual (person / household) level, then we no longer have the data to define geographies on the basis of the same indicators. This is not to say that we couldn't use dwelling counts (from council tax records, for example) and set a threshold for the number for the minimum and maximum to comprise a small area (I have every confidence that David Martin and Samantha Cockings of the University of Southampton can apply their algorithms to construct areas from dwelling locations) but there would be boundary changes in 2021 compared with the 2011 boundaries due to the indicator used, not because population had grown or shrunk over time such that the previous threshold was passed.

Taking a census provides an anchor for other data sources. In addition to the use of census geographies for the dissemination of administrative data, the census informs the collection and dissemination of survey data. The overall sampling fraction and detailed data collection strategy for surveys are often informed by

census data. Sometimes there is over sampling of specific groups (minority ethnic groups) and there can be a plan for respondents from a range of area types whether defined by level of deprivation or by geodemographic classification (yes, both of these schemes can be achieved without using census data but the Index of Multiple Deprivation is released for LSOAs which currently are a census-defined geography, as noted above). Moreover, when released, most large scale surveys include weights to allow for differential non-response and to ensure a representative sample. In the main, these weights will be defined by comparing the survey sample with census data. Additionally, whilst survey microdata may lack geographical detail by way of place names, datasets have area attributes attached to respondent records to indicate the type of place in which people live and these may well be derived from the census. The ONS Longitudinal Study is a prime example.

Postscript

The Beyond 2011 project for England & Wales and equivalents in Scotland and Northern Ireland will be making recommendations soon so we need to respond quickly when asked for evidence which can be used. Please put by half a day to put together a case study and supply it through the appropriate conduit. We will not have a 2021 Census unless we can prove its worth and whilst linked administrative records have incredible potential, the shifts in attitudes needed for this to be achievable are huge. The reality is, we are faced with the gloomy prospect of neither and having to cope with very watered down alternatives unfit for our purposes unless we can help provide information to underpin the business case which will be presented to government.

Further information

Information on Beyond 2011 from our national statistics agencies

ONS: <http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/index.html>

NRS: <http://www.gro-scotland.gov.uk/beyond-2011/>

NISRA: email beyond2011.nisra@dfpni.gov.uk

Thinking about a statistical system

Raymer J, Rees P, Blake A, Boden P, Brown J, Disney G, Lomax N, Norman P & Stillwell J (2012) Conceptual Framework for UK Population and Migration Statistics. Office for National Statistics.
<http://www.personal.leeds.ac.uk/~geopdno/Raymer-et al-2012.pdf>

Devising 2011 Census geographies

Cockings S, Harfoot A, Martin D and Hornby D (2011) Maintaining existing zoning systems using automated zone design techniques: methods for creating the 2011 Census output geographies for England and Wales. *Environment and Planning A* 43(10) 2399 – 2418.
doi:10.1068/a43601

Paul Norman, School of Geography, University of Leeds

Email: p.d.norman@leeds.ac.uk