



# **Beyond 2011**

**Stakeholder Engagement workshop  
Glasgow, 7<sup>th</sup> and 8<sup>th</sup> February 2013**

**Published on 30 April 2013**

## Contents

1.	Introduction .....	3
2.	Current uses of census data .....	3
3.	Alternative data sources used.....	7
4.	Views on alternative sources .....	8
5.	Future census options .....	12
6.	Priorities for alternatives to the census .....	12
7.	Implications of not having the census .....	14
8.	Conclusions .....	14
9.	Our contact details .....	14

## **Beyond 2011 Stakeholder Engagement workshop – 7 & 8 February 2013**

### **1. Introduction**

National Records of Scotland (NRS) has been running a series of stakeholder engagement sessions with representatives from a range of users interested in small area population and socio demographic information. These discussions on the 7<sup>th</sup> and 8<sup>th</sup> February 2013 precede a formal consultation to gather user needs which is being launched in March 2013.

Each session began with a presentation detailing the reasons behind the initiation of the project and the different options that are currently being investigated. This presentation can be found within the [Beyond 2011 Events](#) section on the NRS website. Delegates were invited to discuss how they currently use census information and how any changes could affect their work. The outcome of these discussions is included in sections 2 to 9 of this document.

### **2. Current uses of census data**

Delegates discussed how they currently use the census. The following section details how the census is currently used by the delegates.

#### **2.1 Planning Services**

The majority of delegates predominantly use census data for the planning of services. The delegates use a range of aggregate statistics derived from the census to effectively plan services. Aggregate statistics are a collection of summary counts of the numbers of people, families or households resident in specific geographical areas possessing particular characteristics, or combinations of characteristics drawn from the themes of population, including ethnicity and religion, health, work, and housing. They are derived from analysis of the information provided in returned census forms. Adjustments are made in the outputs for people and households who may have been missed out and not been recorded on a form, or may have been counted more than once by being recorded on more than one form.

A number of the particular characteristics and combinations of characteristics drawn from the census, such as disability and ethnicity, were mentioned by the delegates as being important to the planning of a range of services. A number of delegates emphasised the importance of cross tabulations to planning services. Cross tabulations summarise data based on two or more of the characteristics and can show the interrelationship between two or more characteristics. For example, gender and ethnicity can be cross tabulated with employment information to produce figures on how many unemployed Pakistani women there are within a specified geographical area. Delegates agreed that reliable information on these issues can only be found in the census.

Information on tenure and room numbers is used for housing planning and strategy. It is cross tabulated with information on age and disability to inform housing strategy.

‘We need population and migration statistics to understand community profiles and forecast the housing market.’

‘The census is also important for providing data to support housing strategy.’

A number of the delegates also said that they were looking forward to the release of the 2011 census data and that they would be using data from the new question on disability.

‘We need census data to understand how many people will have needs in terms of learning disabilities and carers. We are very interested in seeing the results from the 2011 Census on this.’

The delegates discussed a number of specific services that require unique estimates from the census. Some delegates use the census for transport planning and emphasised that this is the only source from which they can derive information on land use and transport.

‘The data is used to model the interaction between land use and transport. We supply the outputs to local authorities on travel to work and local transport models. For example, we use census data to create models for the A9. It is essential to the modelling for us and local authorities.’

The [public sector equalities duty](#) (PDF document available on the Press for Change website) arising from the Equalities Act 2010 (on the [legislation.gov.uk](#) website) requires all public bodies to make assessments to ensure they are not discriminating against anyone with a protected characteristic (i.e. age, gender, disability, ethnicity, sexual orientation, religion and belief, etc.). The census provides Local Authorities and NHS Boards with baseline information on these characteristics. Many delegates stressed the importance of the census as the primary source of equality information, particularly relating to ethnicity.

‘The census is the only source of information on ethnicity. We are required to provide information and compare ethnic minority groups.’

‘For us ethnicity information from the census is very important and you can’t get it from other sources.’

## **Geographic coverage**

Aggregate statistics derived from the census are available from output areas to a national level. This flexibility to the statistics was considered by delegates to be very important to the effective planning and targeting of resources. The

delegates emphasised the importance of small area data to their work. Having low level geographical outputs was essential to a number of delegates as it allows them to build up to the different geographical areas and boundaries based on their needs.

‘All boundaries for different services are different in our area. Having data at a low level of geography is how we can determine our information for the different areas. We don’t need data zone level all the time but we do need to split our council.’

‘Having data at a small area level is valuable, especially economic data. Being able to compare employment with areas of deprivation is good.’

One delegate noted that the census is the only source detailed enough to be relevant to the Islands.

‘The census is the only source for information on the islands. All other sources amalgamate the areas and this is not even done consistently.’

The aggregate statistics derived from the census and subsequent Mid -Year Estimates (MYE)<sup>1</sup> are used with other data sources as denominators to calculate service provision. Some delegates use census population estimates as the denominator when estimating prevalence rates (the proportion of the population with an illness or condition at a given time) and the health ratios of health conditions. These estimates are based on both the census as the denominator and other data on the total number with a health condition to provide the numerator. This was important to a number of delegates:

‘The census is essential as it is the denominator to calculate health ratios.’

‘Data from the census gives you prevalence rates.’

Prevalence rates are then used to forecast service needs at the necessary geographical level.

‘We base our [social work services] forecasts on total population numbers and then use prevalence rates to plan our services.’

Population estimates derived from the census are also used as the denominator when calculating outcomes and measuring the effectiveness of service provisions.

‘The census is essential to all work that we do – it’s the denominator for all MYE and outcomes.’

---

**Footnote**

- 1) The mid-year population estimates provide an annual estimate for the population as of 30 June. They are based on census data and records of births, deaths and migration patterns.

A number of delegates mentioned that the census is used in resource allocation calculations:

‘We use the headline figures on population for funding at a local level.’

‘Census data is key to service provision and without that data we won’t be able to make rational resource distribution.’

A number of delegates said that changes to the MYE can have large financial repercussions for local authorities.

‘We need good figures to predict accurately. It affects our grant expenditure.’

## **2.2 Information Requests and Business Planning**

A number of delegates also discussed their role in providing demographic information derived from the census to small businesses to effectively target investment and business planning.

‘The businesses come for planning purposes and profile of the ages to which they will provide services.’

‘There are private enquiries from small businesses. For example, we might get dental practices looking for socio-demographic data.’

## **2.3 Trust in the census**

Across the discussions a number of delegates emphasised that the census is perceived as the most trusted and authoritative source of information. For example:

‘I think the census is a gold standard and it would be hard to replicate it from other sources as no other source has the same quality or reliability.’

‘The census is the single most comprehensive source that we have.’

‘One could question a local survey but not census statistics.’

A number of delegates across the groups also raised the issue that it is the perceived authority and trust in the accuracy of the census that is important to their work. This perceived authority of the census was discussed in the groups as being important to legitimising planning and policy decisions.

‘When plans and policies are developed people object. If we don’t have the data we cannot defend and justify what we are doing.’

‘So much technical information is used to justify what we do. We can’t be criticised if it is evidence-based.’

## 2.4 Limitations to the current census

A number of delegates across the group criticised the frequency of the census and the consequential limited time frame in which it can be used.

‘The census was seen as relevant until about 2007 or 2008 then it wanes a bit.’

‘As years go by we trust less in the census. This is a real problem and it directly affects funding.’

A number of delegates inferred that from their experience of the shelf-life of the 2001 census they would predict the following for the 2011 census.

‘Its 2013 already so we are already 2 years on from when the census was done. It loses its relevance by 2017 and so we only actually get 4 years when it is relevant.’

Some delegates felt that the ten year frequency is not in line with changes in society. For example, some delegates felt that changes in the distance travelled and modes of transport are not captured in the data.

‘Travel habits have changed a lot in the last few years and they are not captured by the data.’

A similar problem was outlined for data on ethnicity, which is also unique to the census at lower geographical outputs.

‘In our small council the ethnicity variability is high and that means that data becomes outdated very soon.’

Many delegates across the groups were concerned that the 10 year frequency missed the population peaks and fluctuations caused by migrant populations.

‘Migration from eastern European countries wasn’t showing up in the Mid -Year Estimates. We could see from our services that translation requests in social work increased massively. This kind of thing makes it difficult to plan translation budgets.’

‘With migration, we can tell who is coming in but not who is going back again. I think there was a peak in 2007 or 2008 which the census will have missed.’

## 3. Alternative data sources used

Delegates were asked about other sources used in their work. A wide range of administrative sources were used including Her Majesty’s Revenue & Customs (HMRC), Department for Work and Pensions (DWP), Community Health Index (CHI), Driver and Vehicle Licensing Agency (DVLA), Electoral

Register, the School Census, the Citizen's Account, NOMIS (official Labour market statistics) and the NHS Central Register (NHSCR).

The delegates also used a variety of surveys including the Scottish House Condition Survey (SHCS), Scottish Household Survey (SHS), Scottish Health Survey (SHeS), the Fuel Poverty Forum, Annual Population Survey (APS), and the Business Register and Employment Survey (BRES).

They also use a range of other sources including management information systems including CACI (Marketing and Information solutions), the Scottish Index of Multiple Deprivation (SIMD), Scottish Neighbourhood Statistics (SNS) and Scottish Longitudinal Study (SLS).

The SIMD was considered to be a very useful tool for community planning by many delegates. Local authorities use SIMD to identify small area concentrations of multiple deprivation. It allows effective targeting of policies and funding. The SIMD ranks small areas (called datazones) from most deprived ranked 1 to least deprived ranked 6,505. The SNS is the Scottish Government's on-going programme to improve the availability, consistency and accessibility of small area statistics in Scotland and was also deemed extremely useful by many delegates. SNS<sup>2</sup> has developed the systems to enable statistics across most policy areas including information about benefits, education, health and the labour market to be brought together on a common small area geography.

'SNS and SIMD are the main drivers for planning in our council area. They are useful for investment regeneration.'

## **4. Views on alternative sources**

### **4.1 Surveys**

Delegates discussed how social surveys could be used as an alternative to the census in their work. The two main surveys discussed were the Scottish Household Survey (SHS) and the Scottish Health Survey (SHeS). Although most delegates agreed that they are very useful sources it was agreed that neither survey had a large enough sample size across all local authorities in Scotland. In order to ensure that results can be said to be representative of the whole population a sufficiently large sample size is required. Usually a sample with between 1,000 to 2,000 interviews will allow for reasonably robust analysis<sup>3</sup>. However the sample size needs to increase if more detailed analysis of sub groups of the population, or aggregate statistics, are required. For example, if analysis is to include results for men and women separately then the sample will be divided in two (leaving 500-1000 per sub-group), but if there are 5 age categories to be analysed the sample will be cut 5 ways making the numbers in each sub-group smaller and the results less robust.

---

#### **Footnotes**

2) More information can be found on the [Scottish Neighbourhood Statistics](#) website.

3) Further information on survey design can be found in the [Survey Methods section](#) of the Scottish Government website.

‘The Scottish Household Survey is good in the sense of the questions it does ask. It is not anywhere near as robust as the census though and the sample size is much too small. However, if the sample size was boosted the information is better and perhaps more akin to what we want.’

‘The beauty of the Scottish Household Survey is cross tabulation and the ability to relate variables, but the sample size is too small to make it worthwhile.’

‘Although the sample size in the Scottish Household Survey is too low, I would not be in favour of removing it. It is a good and useful source.’

There were similar thoughts on the Scottish Health Survey. A number of delegates highlighted the problem that the sample sizes are not large enough for analysis to be robust at lower geographical levels.

‘The Scottish Health Survey is good but all we can do is compare Glasgow to the rest of Scotland and we can’t rely on the sample being good enough in the rest of Scotland.’

‘With the Scottish Health Survey you don’t get any data that is particularly useful unless a boosted sample size is requested.’

## **4.2 Administrative data sources and Data Linkage**

Data Linkage is the joining of two or more administrative or survey datasets to increase the power of analysis possible with the data. [The Data Linkage Framework](#) (available on the Statistics section of the Scottish Government website) aims to address barriers to data linkage, and to widen the range of data linkages that can be carried out, without impinging inappropriately on personal privacy of data subjects. The groups discussed the merits and problems of using administrative data sources and data linkage. For some delegates this was based on their own experience of using administrative data and data linkage projects.

One of the key benefits of data linkage projects discussed was validating local authorities’ own data. For example, a delegate stated:

‘Linking projects are good to validate our own data. When there are discrepancies you can take a note of the gaps. The potential of data linkage projects is exciting and we are happy that is very much in the agenda of the Scottish Government.’

Some delegates discussed work recently completed in North Lanarkshire council to create ‘golden records’, which are administrative records with very high accuracy.

‘Golden records can be created by validating who people are and where they live. There is a council which has achieved these already. It

was done by working with the NHSCR. Councils are now looking in detail at data quality.'

One of the main concerns by delegates was the perceived coverage issues of administrative sources. One delegate stated that some administrative sources can reveal additional information on certain groups that are hard to reach in the current census model.

'We know from our own projects using the School Census data that census data was not completely correct for asylum seekers.'

On the other hand, delegates noted that other administrative data sources, such as the electoral register, might have their own coverage problems and would therefore need to be used alongside other sources.

'The electoral register could be used but there are coverage problems capturing vulnerable groups.'

Problems with the timing of updating or cleaning administrative sources were also discussed in relation to the CHI. The impact of this could potentially be that the population estimates drawn from the CHI could be inflated as patients who have died or moved from a GP's practice could be left on the CHI system. One delegate commented:

'Over-inflation is also a concern. We know that it takes GPs a long time to update records.'

A number of delegates expressed concerns about the comparability of data across different local authorities because of different practices:

'My fear is that you have 32 different local authorities. There is likely to be no consistency in their data. One could have good data quality and one very poor data quality, which makes any comparisons pointless. For this idea to work you would need to make sure data from local authorities is as robust as it is in the census.'

'I think a problem with alternative sources would be that everyone would be off doing their own thing and it would be inconsistent. You would end up with a huge mess.'

#### **4.3 Issues with Data Linkage**

Delegates also had concerns about the practicality, cost and the time that it takes to maintain clean administrative data in order to carry out data linkage:

'Could it be more work to clean up all the datasets than we think? The NHSCR is tidy but it is a full time occupation. Lots of people work on it every day. Could an alternative end up being more expensive than the census?'

One delegate with experience of data linkage also warned of practical problems and potential legal issues:

‘I have various concerns about merging data. Our organisation merged with others and we consequently needed to merge data sets. It’s never as easy as it seems. We had a nightmare deciding which data is the stronger or more reliable. In reality it’s really challenging and there are a lot of technical and legal issues. There are so many regulations and agreements about how the data can be used. That is not to say that it can’t be done and it won’t be better – it’s just really difficult.’

#### **4.4 Privacy and security concerns**

Some delegates were also concerned about the impact that the use of administrative data could have on individual’s privacy.

‘At a personal level I am concerned about a Big Brother society. The public will also have great concerns I’m sure. At least with the census you can choose what to answer.’

One delegate had concerns over the security of data from administrative sources.

‘We do a lot of linking projects which will help us in the future to do analysis. For instance, we are linking home carers with housing data, but these projects find a lot of problems with data security.’

#### **4.5 Perceived trust in alternative sources**

As discussed in [paragraph 2](#) (Current uses of the census), a number of delegates argued that one of the most important aspects of the census is the perceived reliability and accuracy of the census.

‘There is a trust in what is provided, a trust in the census’ consistency and reliability. I think the problem with other data sources is that someone might say ‘I’m not sure about this data set’ and no one will use it. We need to build confidence in alternative data sources.’

‘I think there would certainly need to be trust in the alternatives.’

One delegate raised the issue of this trust being related to branding and its official status and whether or not this is potentially transferable to other sources:

‘Is it possible that we all have more confidence in the census simply because it looks official? If you produced results from administrative sources and they had the NRS stamp on them would people trust them as much as they trust the census?’

A delegate also emphasised the need for transparency and the need for data to be available to the public:

‘I think the wider population need to understand the decisions we make – is it evidence-based or a hunch or prejudice? It needs to be based on information available to all.’

## **5. Future census options**

Some delegates also gave feedback on the traditional census and variant census designs outlined as options in the presentation. A number of delegates were extremely positive about the potential of the option to develop the online completion of the census.

‘I think that developing the online option would be very good. It would allow more detailed questions to be covered for less money.’

There was some uncertainty among the delegates about the short form plus survey option due to the restrictions on the questions applied to the short form and the impact that this could potentially have on the information needed for service provisions. Some delegates said that they would be very keen to see the short form and give feedback on its content.

‘If you were to have the short form/long form method then I think we would need a mock-up of the survey and feedback to you on it.’

## **6. Priorities for alternatives to the census**

The groups were asked to discuss the following aspects of the information they use and their relative importance:

- Quality and Accuracy
- Geographic coverage
- Frequency
- Aggregation

There were mixed opinions on what aspects were most important largely because the uses of the census differed across the delegates. Many delegates were keen to stress the importance of quality and accuracy above frequency.

‘Even though there is a delay I would say it’s better to have more accurate data. Quality is more important. I would rank it above frequency.’

‘Quality and accuracy are the most important aspects.’

However, the frequency of the census was considered a priority for many.

‘Information seems to be out of date. It’s already two years out of date and we can only use it for five years after that. I would prioritise frequency. That would be fantastic.’

‘People are more geographically mobile than they used to be, especially in cities, so more frequent than 10 years would be good.’

Delegates were asked to state what frequency they would like in an ideal scenario. Most thought that having population and socio-demographic data of census quality every five years should be feasible. Some would like data more regularly than this but recognised it would be unrealistic.

‘I would say five years would be the right interval. One year would not allow enough time to do the analysis.’

‘In terms of frequency, we have five year rolling plans, so every five years would be good.’

Geographical coverage was also considered to be important, and delegates recognised this would be especially important for people in rural areas.

‘It’s the geographical spread and sample size across the country that is important.’

‘Being able to get data for small areas is important for the work that we do.’

‘Geographical coverage is also important as we need to pinpoint specific towns with problems, especially with depopulation.’

Aggregation was also important to some delegates, where individual level records were seen as valuable for data matching and cross tabulation.

‘I really value the fact that we can do cross tabulations. We don’t have any other sources that can do this.’

‘All of these things are important. Comprehensive cross tabulation is the most important aspect. Coverage in terms of attributes should be up there.’

For many delegates, however, aggregation or disaggregation was not more important than the other factors.

‘In our case we’re not interested in individual level records, so we think accuracy, frequency and geography are most important.’

‘I would give up some disaggregate information if the frequency was better.’

## **7. Implications of not having the census**

The groups discussed what the main implications would be if the census outputs were not available.

‘There would be a panic if we didn’t have the census.’

For many delegates the census is essential to service provision and to understanding their performance. Some delegates expressed concern that if they did not have the census they would ultimately need to produce their own census-like event for adequate coverage at lower geographic levels, which would have large financial implications.

‘Without the census we would not know what service provisions are needed. We would not know how well we were performing. We would have to reinvent the census at lower levels.’

‘There would be huge cost implications as we need this information, and we are used to having this information at our fingertips.’

## **8. Conclusions**

The discussions generated were useful and allowed the Beyond 2011 team to see how the census is used across Scotland. Delegates emphasised the importance of the census to their work. There was also a lot of discussion on the issues surrounding data linkage and the potential solutions to these problems.

The information included in this report will be added to the bank of information currently being collected by the Beyond 2011 team and will be used in the evaluation of any census alternatives.

Delegates will also be kept informed about any future events relating to Census options and will be sent a link to the formal consultation in early 2013.

## **9. Our contact details**

For further information or enquiries about the Beyond 2011 please contact:

Beyond 2011  
National Records of Scotland,  
Ladywell House  
Edinburgh  
EH12 7TF  
Tel: 0131 314 4299  
Email: [beyond2011@gro-scotland.gsi.gov.uk](mailto:beyond2011@gro-scotland.gsi.gov.uk)