

## Draft Paper For Ukcdmac Meeting On 6 October 2010

### **Methodological Options for Moving From A Six-Month Enumeration Base To A 12 Months Output Base For Scotland**

**UK Statistics Authority (UKSA) Census Assessment Phase 1, recommendation 1 requires the Census Offices to:**

Publish explanations of how the census estimates relating to the resident population, based on 12-month residency, will be derived from the information collected about three and six-month residency.'

**The agreed Census Offices response, sent to the UKSA on 30 July, states:**

An explanation of the different population bases to be used for enumeration and outputs in the 2011 Census in England, Wales and Northern Ireland is outlined below. It also explains the rationale for each, and the process agreed for moving from enumeration to outputs.

**Note:** A separate paper, outlining the equivalent bases and process for Scotland, is being prepared by General Register Office for Scotland (GROS)  
(my emphasis).

**This paper sets out some suggestions on the methodological approach to be taken in outlining the equivalent bases and process for Scotland.**

#### Response to Requirement 1 for England, Wales and Northern Ireland

This is appended as [Annex A](#). It explains, at 2.1.1, that:

"further consultation identified increasing requirements for information on short-term migrants and it was agreed that the census enumeration base would be expanded to meet this need. To achieve this, it was agreed that a full census return would be collected from everyone who has been, or intends to be, resident in the United Kingdom (UK) for a period of three months or more in the 2011 Census."

But with a footnote adding:

The Census in England, Wales and Northern Ireland will be expanded to collect this information, but the same strength of requirement has not been identified in Scotland.

At 2.1.3 the response presents an algorithm, in diagrammatic form, showing how the Office for National Statistics (ONS) propose to identify usual residents for the main output base viz. those people who have been, or intend to be, a resident of the UK for a period of 12 months or more, using the Census variables country of birth, month/year of arrival, and intended length of stay.

## Draft Paper For Ukcdmac Meeting On 6 October 2010

### Discussion

The task for GROS is to identify how best to put together a process which does the equivalent job for Scottish data in the absence of an intended length of stay variable. (The decision by ONS to use the intended length of stay variable in this way came after GROS had analysed user needs for data from the Census in Scotland and concluded that there was little demand in Scotland for information on intended length of stay, and had dropped it and related items from the draft questionnaire to make space for other variables of more interest to Scottish users.)

So we want to know for those who arrived in the UK on or after April 2010, what proportion of them would have described their intended total length of stay as more or less than 12 months, if an intended length of stay question had been asked in the Census in Scotland.

We will primarily look at this by using the ONS Census rehearsal dataset. We will look at how the relation between X, where

X = % of those who arrived <1 year ago who replied to the intention to stay question, who replied that they intended to stay for a year or more varied against other relevant census variables, including:

- Age
- Marital status
- Household Size
- Economic Activity Status
- Country of birth

The intention is to use this analysis to develop a decision rule to be used to determine which of those who completed the Census in Scotland and stated that they had arrived < 1 year ago should be included in the census file as usual residents.

Once such a decision rule has been developed it could be applied deterministically, with the inclusion/non-inclusion of each case determined by direct application of the rule to the details of the individual case, or probabilistically – in the latter instance we would derive from the ONS rehearsal data the probability that cases with each combination of the values of a small list of key variables would report intended length of stay as a year or more, and then apply a random process to include the relevant respondents in the dataset with the appropriate probability.

An obvious weak point in the method is that the mix of recent arrivals in England and Wales at the rehearsal could be very different than the mix in Scotland in the Census proper. Accordingly we will review the data that we have available from other sources on immigrants from overseas to Scotland and England & Wales, to help inform the decision whether, once we have derived a decision rule from the ONS rehearsal data we will need to tweak it for the Scottish situation.

## Draft Paper For Ukcdmac Meeting On 6 October 2010

The discussion above concerns decisions on the inclusion/non-inclusion of individuals in the census file according to various items of data. Of course in practice it may be the household, rather than the individual, which we have to decide whether to include in the Census datafile. GROS would expect, in the application of the rules developed in this area, to align with whatever treatment of this aspect that ONS adopt.

**UKCDMAC comments and suggestions on these proposals are invited.**

**GROS September 2010**

## Draft Paper For Ukcdmac Meeting On 6 October 2010

### ANNEX A

#### Response to Requirement 1 for England, Wales and Northern Ireland

##### **2.1. Initial agreement on population base**

The three UK statistical offices agreed in 2005 that the population base to be used for enumeration in the 2011 Census should be Usual Residents and Visitors. This was consistent with the Registrars General agreement to harmonise the three UK Censuses as closely as possible.

This decision was reached after extensive research and consultation, summarised in the paper: 'Selection of the population base for the 2011 Census enumeration'.

At that time, the exact definition of a usual resident to be used had not been finalised, as this was planned for agreement after further consultation and extensive questionnaire testing.

##### **2.1.1. Population base for enumeration**

At the time of the 2007 Census Test, it was assumed that the enumeration base for the 2011 Census would be broadly consistent with that used in 2001, but with the addition of visitors, and the questionnaire reflected this.

However, further consultation identified increasing requirements for information on short-term migrants and it was agreed that the census enumeration base would be expanded to meet this need. To achieve this, it was agreed that a full census return would be collected from everyone who has been, or intends to be, resident in the UK for a period of three months or more in the 2011 Census.

This enumeration base has been evaluated through qualitative cognitive testing and quantitative postal tests and ONS is confident that it will collect the required information in the 2011 Census.

##### **2.1.2. Main population base for outputs**

After reviewing the 2001 Census and a series of discussions between census and demography staff across all three UK statistical offices, it has been agreed that the main output base for the 2011 Census should be 'Usual Residents'. It has also been agreed that usual residents should be defined as those people who have been, or intend to be a resident of the UK for a period of 12 months or more.

This population base will be used for all main outputs from the 2011 Census, across the whole of the UK, and the process for identifying the usual residents in the outputs database in England, Wales and Northern Ireland is outlined in section 5. In addition to this, a subset of outputs (yet to be defined) will be produced for the short-term (not usually resident; in the UK for 3 – 12 months) population.

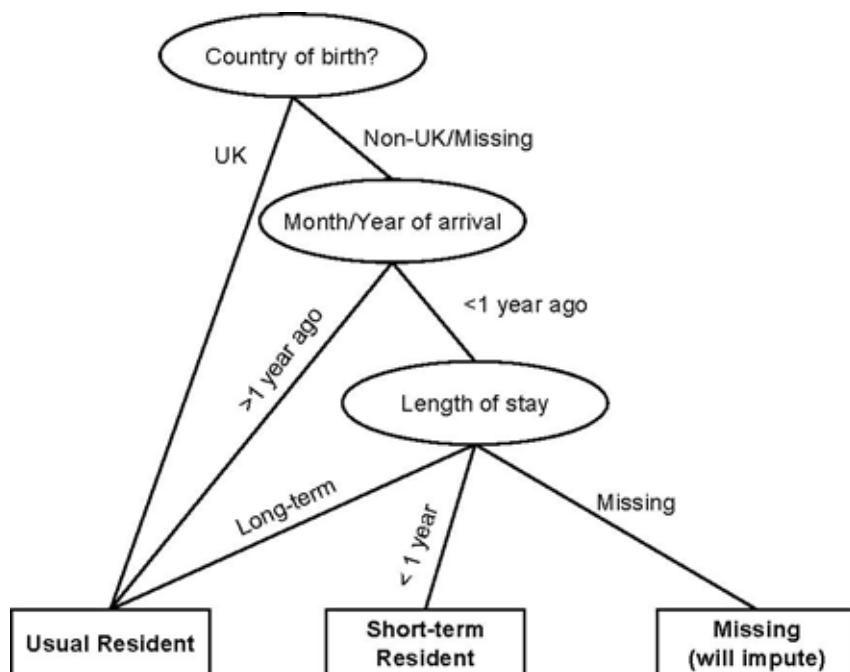
## Draft Paper For Ukcdmac Meeting On 6 October 2010

### 2.1.3. Identifying usual residents

To make it possible to distinguish between usual residents and short-term migrants in the census database when producing outputs, ONS has developed and tested a question on 'Intended length of stay in the UK'.

Respondents who have lived in the UK for less than 12 months will be asked to indicate whether their overall stay is likely to be 3 less than 6 months, less than 12 months or long-term.

This information is used, together with responses to 'Country of Birth and Month and Year of Arrival' to identify all usual residents, as shown below:



All people born in the UK are assumed to be usually resident although, in reality, a very small proportion will not. This assumption will avoid confusion for the majority of respondents.

Missing responses to these questions will be imputed, using the same methodology as for the rest of the data, based on respondents with similar characteristics. This will enable these people to also be categorised as either a usual or short-term resident.

### 2.1.4. Risk of using intention to stay question

For all questions included in the census for the first time, there is a risk that it will be unacceptable or confusing to respondents, or that the data collected will be inaccurate or incomplete. These risks are perceived as particularly high for the Intention to stay question, as it is a sensitive subject. In addition, the information identifies those who should be included in the main output base and those who should be included in the alternate population base.

As with all questions developed for the census, this question has been thoroughly tested, and ONS is confident that it is acceptable to, and

## Draft Paper For Ukcdmac Meeting On 6 October 2010

understood by respondents, who were able to answer the question confidently and accurately.

However, when the question was included in the 2009 Census Rehearsal in England and Wales, a far higher level of item non-response (24.5 per cent) was experienced than anticipated. As non-response for other questions was significantly lower, this was a cause for concern and further analysis was undertaken.

This analysis showed that 68 per cent of these non-responders were students, almost half of whom had arrived in the UK within the last month. It is considered probable that, as the rehearsal was held in October at the start of the college/university term (whereas the census will be held when this cohort will have been in the UK for six months), the level of item non-response among them will be significantly reduced. If students are excluded, the level of item non-response to this question is comparable to that for others, and lower than for several.

Of the remaining non-respondents to this question, 6.2 per cent (of all non-respondents) had been in the UK for more than 11 months, and their non-response is likely to be because they consider themselves usually resident.

More generally, respondents not answering this question had usually answered others sufficiently accurately (over 90 per cent of non-student adults had answered the labour market questions) to enable imputation with confidence.

After considering the analysis summarised above, together with the different dates for census and rehearsal (which is likely to significantly reduce the number of recent arrivals in the UK), ONS is confident that this information will be of sufficient quality to allow the accurate differentiation between usual and short-term residents in the 2011 Census. To help ensure this, significant extra guidance will be available, through the census website and contact centre, to assist recent immigrants in general, and students in particular.