

Samples of Anonymised Records from the 1991 and 2001 Censuses

Where things are and where they
are going

Prof. Angela Dale

CCSR, University of Manchester

Re-cap on 1991 SARs

- 1% Household sample:
 - Hierarchical structure, all people in household
 - No individual records where 12+ pp in hhold
 - Limited to regional geography
 - Extensive individual detail, eg:
 - SOC (358 categories); industry; age, ethnic group
- 2% Individual sample:
 - Flat structure, less detail, more geography
 - Large LAs (120K threshold); Minor SOC (77)

2001 SARs: request to ONS

1% Household SAR - similar to 1991

3% Individual SAR - smaller population threshold, 90K

JRSS (A) paper assessing confidentiality

Additional file: Small Area Microdata:

- to provide more geography (sub LA) with less individual details

2001 SARs: response from ONS

1% Household SAR

- concern over level of individual detail especially for large households
 - No occupation below minor SOC
 - Large households – age grouping at 7+
 - No information at 10+
- Concerns over households of mixed ethnic groups (where mix does not include white)



3% Individual file

Accepted increase in sample to 3%

But concerns of detail versus geography

- File will maximise on individual detail and lose geography
 - only region retained
 - but full ethnic group, religion, SOC minor, NS-SEC, all other census variables
 - LA information to be imputed by Paul Williamson



Small Area Microdata

- Request accepted in principle
- 5% file at LA level of geography
- Much less detail than 3% Individual file
- More work to be done on exact specification

Timetable: Individual SAR

- Specification for 3% Individual SAR finalised (almost)
- Data checks for 'special uniques', started Monday June 16th
- Work needed to resolve unusual cases that may be risky
- Autumn target date for data from ONS
 - Release to research community – December 2003 ??

Timetable: Household SAR

- Some issues still unresolved:
 - Large households – we plan to impute back the records for large households using LFS
 - Households with mixed-ethnicity
- This will be next to be finalised
 - Aim to get agreement in next weeks



Scotland

- Differences caused by devolution
- Question over SARs for Scotland
 - Scotland released more risky census tables
 - makes it harder to ensure safety of microdata
- Scottish data likely to be much less detailed



The good news

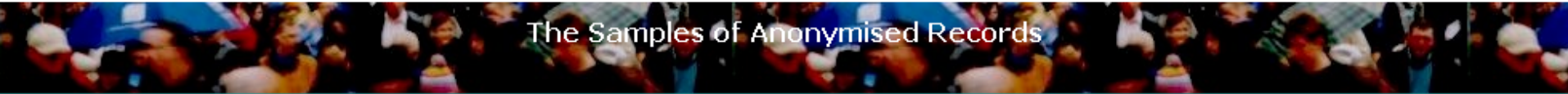
- ONS have done additional occupational coding for individuals not coded by census
 - Last worked >5<10 years ago (16-65)
 - 65-74s not currently working who worked in the previous 10 years
- Have accepted in principle release of an additional file



Details of SARs

SARs web-site: www.ccsr.ac.uk/sars

- Contains specification of Individual SAR
- Discussion over confidentiality
- Information on web registration and access:
 - now in place for 1991 SARs



The Samples of Anonymised Records

[Home](#)

[What are SARs?](#) | [Getting Data](#) | [2001 SARs](#) | [1991 Documentation](#) | [Support & Training](#) | [FAQs](#)

[2001 progress](#)

Welcome to the Samples of Anonymised Records (SARs) at the Cathie Marsh Centre for Census and Survey Research ([CCSR](#)). The Centre [supports](#) and [disseminates](#) the 1991 and forthcoming [2001 Samples of Anonymised Records](#).

[News](#)

The SARs are samples of individual Census records which have been adapted to preserve respondents' confidentiality. This was achieved by removing identifying information and restricting the amount of detail available. The result are very large datasets resembling survey data. They containing a wealth of information for social scientists and policy makers which has been widely used since the release of the 1991 SARs in 1993. They are fully described in the '[What are the SARs?](#)' section of the site.

[Use](#)

The most recent Census took place on April 29th 2001. Results will be available in 2003 including a series of new SAR files. CCSR are currently preparing for the release of these data, and plans are available in the pages for the [2001 SARs](#) including the near final specification for the Individual SAR.

[Access](#)

Meanwhile, the 1991 SARs are still available to registered users.

[Support](#)

This site contains:

[Teaching](#)

- Details on how to [register](#), [obtain](#) and [use](#) the 1991 and 2001 data.
- Information on user [support](#) and [training](#).
- Detailed documentation.
- [Suggested uses](#) for the data and a [review of research already undertaken](#) using the 1991 data.
- The [specification](#) and [timetable](#) for the 2001 SARs

[Training](#)

[Findings](#)

[Registration](#)

Email queries should be sent to sars-helpdesk@man.ac.uk to ensure the fastest response. The SARs helpdesk can also be reached by telephone on +44 (0) 161 275 4735. There is also a [SAR mailing list](#).

[Publications](#)

[Events](#)

[Search](#)

[Contact SARs Support](#) | [SARs team](#) | [CCSR](#)



Access for 2001 SARs

- Academic access via web-based registration and web download
- Non-academic access
 - Via CDROM or web with special password
- Format of files:
 - SPSS portable file or STATA
 - Files with analysis software (NSDstat)

Costs for 2001 SARs

- Costs reflect added value – not data
- Business - £1000 for each SAR
- LAs, HAs, Central Government
 - discounted rate of £500 per file
- data available to 10 named people
 - One person as representative
 - All ten users receive support and a free place on a training course.