

SARs NEWSLETTER

August 2006

Samples of Anonymised Records from the 1991 and 2001 Census

CCSR, School of Social Sciences, University of Manchester, Manchester, M13 9PL
www.ccsr.ac.uk/sars/ sars-helpdesk@manchester.ac.uk 0161 275 4735

The SARs for 2001

Five different microdata files are now available:

- 3% Individual SAR (Licensed)
- 3% Individual Controlled Access Microdata Sample
- 1% Special Licence Household SAR
- 1% Household Controlled Access Microdata Sample
- 5% Small Area Microdata

A SARs User Meeting on 16 November 2006 will present results from analysis and look forward to the 2011 SARs (see page 7)

CONTENTS

Editorial.....	2
The data files.....	2
Access to the SARs.....	3
Programme of development and dissemination from August 2006.....	5
Seminars and workshops.....	5
SARs introductory workshop: focus on the Small Area Microdata	6
CCSR short courses.....	6
SARs user meeting.....	7
Contact details	8
The new SARs team (August 2006-August 2011).....	8
Resources for the SARs.....	8
Publications.....	8

Editorial

The Newsletter comes at the end of the five-year funding period for the Census Microdata Unit at CCSR, University of Manchester, led by Angela Dale and the beginning of a new five-year period under the leadership of Ed Fieldhouse. It also comes at the point in time when all the 2001 microdata files are now available and documented. We have therefore set out exactly what files are available and how they can be accessed.

Although the process of negotiating the files was long and difficult the final result is much better than we could have anticipated. As the paragraphs below spell out, we now have more data available than in 1991 and, although less detail is available through the Standard End User Licence, much more detailed information is available in controlled settings.

We would like to take this opportunity of thanking so many people for their advice, help and support in achieving this outcome. In particular we are grateful to the many users who provided their carefully considered views about particular variables and file content. We are also very grateful to the ONS SARs Project Board which met every six weeks from December 2003 to January 2006 and included representatives from all relevant sections of ONS, GROS and NISRA.

The data files

In total, five different microdata files are now available with varying amounts of detail and three different modes of access, with the level of security directly related to the level of detail in the data. More details on all these files are available at www.ccsr.ac.uk/sars.

1. *The Individual SAR (Licensed)*

A 3% sample of individuals from all countries of the UK, comprising 1.84 million people. It covers the full range of Census topics – a total of 88 variables, including some derived variables to represent summary household variables such as the number of employed adults in the household. In this file the lowest level of geography is Government Office Region. A second edition of the data was supplied to CCSR in 2005 with improved detail on ethnicity and religion. Version 2 was published on 1 April 2005. Access to the Individual SAR uses the standard End User Licence via Athens operated by the Census Registration Unit. The data are also available to non-academics and usually supplied on CD.

2. *The Individual CAMS*

ONS have made available a much more detailed version of the *same* sample – a Controlled Access Microdata Sample (CAMS). This file has the full detail recorded on the census database for most variables (e.g. over 200 country of birth categories); geography at the level of local authority and a large number of additional derived variables. For example, the Index of Multiple Deprivation scores are available for the Super Output Area of enumeration. However, conditions of access are much more rigorous than for the Individual SAR. The file can only be accessed within one of four ONS offices in England and Wales. There is a rigorous application procedure implemented by ONS and careful vetting of outputs by ONS. However, a codebook describing the file contents is available from the CCSR SARs website.

ONS has made available a test version of this file (with a smaller sample and heavy perturbation) with a standard EUL licence.

3. The Special Licence Household SAR

The third file is a 1% sample of households and all individuals in the households, for England and Wales only. Again, all census topics are represented. However, this file gave rise to much greater concerns over confidentiality than the Individual file. This was because information for all individuals in each household could be linked and quickly made the household unique in the population using only a few variables such as age, sex, occupation and educational level. These concerns resulted in the file containing no geography at all and being available only for England and Wales (different questions used in Northern Ireland and Scotland meant that if these countries were included they could always be identified and their small population size increased the chances of uniqueness). A codebook describing the file contents is available from the CCSR SARs website.

Access to this file is via a 'Special Licence', operated by ONS, which imposes more restrictions than the standard End User Licence. This is administered by UKDA and is also used for some of the government surveys to allow access to more detail.

4. The Household CAMS

The SL-Household SAR is also available as a Controlled Access file in exactly the same way as the Individual CAMS file. As with the Individual file a large number of derived variables have been added to the file. However, this tightly controlled file also contains data for Scotland and NI. As in the Individual CAMS, the lowest level of geography is the local authority. A codebook describing the file contents is available from the CCSR SARs website.

5. The Small Area Microdata

Finally, the Small Area Microdata (SAM) represents an innovation for the 2001 census and for microdata in general. It is a 5% sample of individuals from across the UK, comprising 2.96 million records. The file is designed to provide the lowest geographical level that could be achieved with microdata distributed using the standard EUL via the census registration service. Geography is at the level of local authority (parliamentary constituency in NI) and, to offset this level of geographical detail, the level of detail on other variables is restricted. It is accessed using the standard EUL.

Together, these files provide a rich source of census data and the SAM, in particular, provides a welcome bridge between the pre-defined census tables and the detailed individual microdata with little geography. One of the key aims of the next phase of the programme will be to ensure the innovative use of these data files.

Access to the SARs

The SARs files are downloadable from the SARs website for users who have registered with CRS using the standard EUL except in the following circumstances:

- where they are non-academic users
- where the level of detail in the data have resulted in more restrictive conditions of access

The Individual SAR and Small Area Microdata

The Individual SAR and Small Area Microdata are both available under standard EUL conditions from the SARs website. Users can select files in the following formats:

- SPSS save files
- SPSS portable files
- Stata files
- Tab delimited text files

In addition, users can select a file which contains variables which indicate items that have been imputed, either by the One Number Census process or by the perturbation conducted on the file (imputed items are not flagged in the SAM for Scotland or Northern Ireland). Derived Variables are available in SPSS portable format and can be merged into the main file.

We encourage users to begin by exploring the data online using Nesstar. Nesstar is a data exploration and analysis tool that allows users to look at univariate frequencies and variable descriptions without needing to register for the data. Registered users can use Nesstar to undertake more complex analyses and to download subsets of data in a wide variety of formats.

Any users without access to a standard statistical package can download the SARs together with an analysis package, NSDstat, free of charge. NSDstat is a user-friendly desktop statistical analysis package with the power to process large amounts of data in minimal time. Features of NSDstat include univariate, bivariate and multivariate analysis, graphical representation of data, regression, weighting and recoding. This forms an important part of our strategy to widen use of the data to students in Further Education.

For UK non-academic users wishing to use these files, we supply data on a CD-ROM. We have an agreement with ONS to supply these files to overseas users where it is clear that the application is from a *bona fide* organisation and the SARs will be used for research. ONS consider each application following a recommendation by CCSR.

SL-Household SAR

The Household SAR is available under a Special Licence that is administered by the UK Data Archive. This imposes more restrictions on the way in which the data are stored and transported. For example, the file must be used on a PC in an office of the employing organization. If the PC is connected to the internet it must be through a secure server rather than a dial-up line. Data cannot be used from home or moved around on a laptop. The Household SAR can be used for research purposes by any organisation (not just academic users). It cannot, however, be sent outside the UK. These additional safeguards are to protect confidentiality.

Controlled Access Microdata Samples (CAMS)

As described above, access to the CAMS is in a safe-setting in one of four ONS sites: London, Titchfield, Newport and Southport. The data files are held in a Virtual Microdata Lab (VML). The VML is a server at Drummond Gate to which users have access remotely from dumb terminals. Each user has an allocated space on the server in which they can work using SPSS, STATA or MLWin.

The user interacts directly with the data; syntax files can be prepared beforehand and read onto the server before the user arrives. Outputs are stored on a secure separate area of the server and are emailed to the user after screening by ONS staff.

Applications for the CAMS are handled by ONS. The application forms, a description of the application process, the criteria used to assess applications and the confidentiality guidelines are available on the ONS website at www.statistics.gov.uk/census2001/sar_cams.asp. All applications are considered by Census Research Access Board on which the CCSR SARs team sits. The application process takes 1-2 months. Only non-disclosive results are allowed out of ONS. There is a charge of £52 per day + VAT for use of the VML.

To facilitate the preparation of syntax files, a heavily perturbed subsample of the Individual CAMS is available to enable CAMS users to test syntax. The file was produced by ONS with the assistance of CCSR and is available from CCSR under an End User Licence. This file maintains univariate but not multivariate distributions and contains nearly all of Individual CAM variables. Please note: This dummy dataset is not suitable for research purposes.

Websites

The SARs website at CCSR: www.ccsr.ac.uk/sars

Census Registration Service: <http://census.data-archive.ac.uk>

Controlled Access Microdata files:
www.statistics.gov.uk/census2001/sar_CAMS.asp

Programme of development and dissemination from August 2006

From 1 August 2006, the SARs are supported by a new round of funding under the ESRC Census programme (see www.census.ac.uk). Under the new programme the SARs team will be lead by Professor Ed Fieldhouse who has been part of the SARs team since 1993. We are delighted to retain all members of the existing SARS team which has served the SARs user community over the past five years. We are also pleased to welcome a new member of the team, Anthony Rafferty (details of the team is provided below). With the data now delivered, under the new programme there will be an emphasis on providing continuing support for users; providing added value to the SAR data through the production of derived variables, documentation and training materials; contributing to an integrated census data service through collaboration with other census units; and planning ahead for microdata from the 2011 Census.

Seminars and workshops

The SARs team is very happy to come and run a workshop or give a seminar at your home institutions. Please contact Gill Meadows in the first instance at gillian.meadows@manchester.ac.uk.

For details of all forthcoming SARs events please go to:
<http://www.ccsr.ac.uk/sars/events/>

SARs introductory workshop: focus on the Small Area Microdata

13 October 2006
CCSR, University of Manchester

This workshop will provide information on the various SARs data files that are available from the 2001 Census. There will be a particular focus on the Small Area Microdata (SAM) which contains local authority level identifiers for sampled individuals, thus providing better geographical referencing than other licensed SAR data. The afternoon session will include a demonstration on mapping census data and a practical and there will be plenty of opportunity for questions and discussion.

It is possible to book for the morning only. For more details and to book a place please go to <http://www.ccsr.ac.uk/sars/events/2006-10-13/>

Programme

10.45	Coffee and registration
11.00	Introduction to the SARs
11.45	The Small Area Microdata: content and access
12.15	Research potential of the SAM
1.00	Lunch
2.00	Mapping census data
2.30	Hands on session: using the SAM
3.45	Questions and discussion
4.00	End

CCSR short courses

CCSR run a programme of short courses on analysis methods relevant to users of the SARs. The next courses to run include Introduction to Data Mining, Starting SPSS and Introduction to Data Analysis. Courses are offered at introductory, intermediate and advanced level. The complete list of short courses is available at www.ccsr.ac.uk/courses/.

Bursaries are available for all courses but are limited and awarded on a first-come, first-served basis. Details can be found at www.ccsr.ac.uk/courses/bursary.html.

September to December 2006

18 September:	Introduction to Data Mining
20 September:	Starting SPSS
21 September:	Introduction to Data Analysis Part 1
22 September:	Introduction to Data Analysis Part 2
25-26 September:	Causal Analysis in STATA
27-29 September:	Generalised Latent, Linear & Mixed Models
27 October:	Applied Statistics for Labour Research
20 – 24 November:	Spatial Regression Analysis
8 December:	Introduction to Stata (Mixed Course)
18 December:	Qualitative Comparative Analysis

For further information about these or any other CCSR courses, please contact Katey Matthews on katey.matthews@manchester.ac.uk or 0161 275 4736 from Monday to Wednesday, 9.00am – 4.30pm.

SARs user meeting



16 November 2006
Royal Statistical Society, London

Draft programme

Please note that this programme is subject to confirmation and may therefore change. A final version will be available from the SARs website at <http://www.ccsr.ac.uk/sars/events/2006-11-16/>

- 9.30 Coffee
- 9.55 Welcome and introduction
- 10.00 Changing Patterns of Ethnic Minority Self-Employment in the UK:
Evidence from Census Microdata
*Ken Clark, Economics, University of Manchester and Stephen
Drinkwater, University of Surrey*
- 10.30 Are ethnic penalties in the labour market reduced for the generations
born in the UK?
Ludi Simpson, University of Manchester
- 11.00 Employment patterns amongst ethnic minority women
Jo Wathan, CCSR, University of Manchester
- 11.30 Coffee
- 12.00 ONS plans for the 2011 Census and implications for microdata
Pete Benton and Ian Cope, ONS
- 12.45 Research potential of the Small Area Microdata
Tony Champion, CURDS, University of Newcastle
- 1.15 Lunch
- 2.00 Why are boy babies less religious than girl babies?
David Voas, CCSR, University of Manchester
- 2.30 Statistical Disclosure Control Policy and Approach for 2011 Census
outputs
Jane Longhurst and Angele Storey, ONS
- 3.15 Panel discussion on requirements for SARs for 2011
Panel members: Simon King, Keith Spicer, Jane Longhurst, Angele
Storey (ONS), Ed Fieldhouse, Jo Wathan (CCSR, University of
Manchester)
- 4.00 End

SARs support at CCSR

Contact details

Cathie Marsh Centre for Census and Survey Research (CCSR)
School of Social Science
University of Manchester
Crawford House
Manchester
M13 9PL

If you have any queries or comments about the SARs please direct them to the SARs helpdesk by one of the methods below. The helpdesk is staffed from 9:00am to 5:00pm Monday to Friday.

Email: sars-helpdesk@manchester.ac.uk
Helpline: +44 (0) 161 275 4735
Fax: +44 (0) 161 275 4722
Website: www.ccsr.ac.uk/sars

The new SARs team (August 2006-August 2011)

Ed Fieldhouse (25%)	Director
Jo Wathan (50%)	Associate director: User support and development
Sam Smith (50%)	Associate director: IT
Reza Afkhami (50%)	Research support
Anthony Rafferty (25%)	Research support – until end 2007
Gill Meadows (50%)	Administration
Angela Dale (5%)	Advisory

Resources for the SARs

The SARs website contains important resources to facilitate your analysis of the data. These include:

- Frequently Asked Questions
- Getting Started with the SARs
- User Guide to the SARs - essential documentation for all users of the SARs
- Key findings from the 1991 SARs
- Introductions to common software packages
 - Analysing the SARs in SPSS
 - Analysing the SARs with STATA

Publications

A list of SAR-related publications and past copies of SARs newsletters are available at www.ccsr.ac.uk/sars/publications/. To submit a publication based on research conducted with the SARs you can email gillian.meadows@manchester.ac.uk or use the web form at www.ccsr.ac.uk/sars/publications/submit.html.

We are currently carrying out a publication trawl. If you have any publications based on the SARs please help us to keep our list up-to-date by responding to the letter.