

SARs Newsletter

Samples of Anonymised Records from the 1991 and 2001 Census

Report of the International Census Conference

Census Microdata: findings and futures

1 – 3 September 2008
University of Manchester

The world map at the registration desk of the Census microdata conference was visible evidence of the global spread of microdata sources. Whether the data are stand alone national resources or part of the wider IPUMS, census microdata from an increasing number of countries are becoming available and, as demonstrated at this conference, being used in valuable and innovative research. Presentations drew on microdata sources from every continent.



Denise Lievesley, kick started the event with a broad international overview on census microdata from the perspective of an official statistician. Denise, whose previous roles have included special advisor to the African Centre for Statistics for the UN and Head of Statistics for UNESCO, talked about the extrinsic value of data, which comes from its importance in providing an evidence base for good decision making. Setting the scene for later presentations, she noted cultural differences in attitudes to census taking and data dissemination.

Such differences were very much in evidence when presenters from Australia, the UK, Germany, France, Canada and Denmark reflected on the processes involved in creating microdata and making microdata samples accessible. A variety of means of producing census-like outputs are being undertaken whether that is a traditional census (UK, Canada), a rolling census (France), survey approaches (Germany) or linked administrative data (Denmark).

Data access decisions are driven by common concerns of balancing utility and confidentiality and are increasingly informed by national and international statistics law. However, solutions vary by country between public or licensed microdata files, safe settings and remote access with many countries drawing on multiple modes to satisfy different user types.



The justification for the work involved in creating these files was readily evidenced in the wide range of research papers presented. Papers covered a range of topics including migration, health, employment, family formation and poverty mapping. Research not only used single census national datasets but were used alongside other data sources, as part of longitudinal datasets and in comparative research.

The conference was a vibrant celebration of the high quality research that has been achieved with census microdata on a world stage, and the issues that will shape census microdata to come. The organisers are very grateful to presenters and chairs for their contribution to the event, and for making their abstracts and slides available online for all to read. These are available at www.ccsr.ac.uk/sars/conference/

Contents

Routes to access: what affects the decisions that ONS takes about research data access.....	2
Building a business case for SARs microdata from the 2011 Census.....	3
Publications using the Samples of Anonymised Records.....	3
The Household Hierarchical Data.....	3
Census Portal.....	4
Forthcoming events.....	4
Recent publications.....	4

Paul J. Jackson, Office for National Statistics

Introduction

Decisions on data access are made within a number of related frameworks (legal, government policy, statistics and research strategy, etc). Each has its special characteristics (rate of change, relative influence, etc), each has its different effect on decision making and each can be modified by successful lobbying by the stakeholders for research access to data. The presentation at the International Census Conference looked at the current state of frameworks for decision making for research access to data, the prospects for change, and how to achieve change where it is needed.

Legal framework

The legal framework for research access to ONS data has never been better than it is today.

The Statistics and Registration Service Act came into effect on 1 April 2008 and now provides a flexible and powerful statutory framework for ONS to enable research access to data. It does well the two things most necessary for research access – it makes the wrongful use of personal data a crime, and opens a gateway determined by the fitness and properness of the researcher. Correctly, in my view, the focus is on the suitability of the researcher and the research, not the design of the data. To future-proof the benefits for research data access, our use of this law must stand up to the scrutiny of Judicial Review. It is the responsibility of ONS and the research community to now build practices that operate legally and conform with the public's expectations. If we fail, we may revert to a less satisfactory situation where only non-disclosive data is available for research.

Government policies

The Government has a positive attitude to sharing data and ministers are determined to improve services by sharing data to improve the evidence base for their decisions.

However, further data losses from any Department or the disclosure of ONS confidential data would be extremely damaging to the public acceptability of research access to confidential data. The Review of Data Handling in Government is the reaction, and ONS access arrangements must be compliant with its. The continuing support, and patience, of the research community is vital. If the research community wants to hold confidential government data, it must make the same necessary changes to organisational and technical standards for data handling that ONS and other departments have to make. It is no less than the public deserve.

National strategies

ESRC's National Data Strategy provides a clear direction for priority research topics. ONS has very similar priorities in its statistical programme.

Priorities are just that – they are meant to be used, and will be used, to prefer support for some research projects over others if and when resources are limited. The NDS addresses infrastructures too. The UK is on the verge of a world-leading collection of data archives, remote access services, and data laboratories for research access. They need to be linked and placed in a spectrum as a joined-up service (albeit a service with many stakeholders), and made easier to use.

Organisation and investment

ONS has a challenging financial settlement. Evidence based policy making is priority, and resources are tight.

Data handling and risk management are becoming ever more expensive, whether done by the design of 'safe data' or the distribution of confidential data. Efficiency gain, doing more with less, is a major driver for change in ONS data access. Great efficiency gains can be realised when one source is available to many users in a flexible, but safe, environment. Therefore ONS needs now more than ever the help of ESDS to find acceptable solutions for data access with these efficiency gains - the Secure Data Service, for example, ought to be one of the innovations that bring these advantages to the research community and ONS.

Partners, Relationships and Trust

The situation is very positive for research data access.

ONS has an excellent relationship with ESDS/ESRC and researchers at high level. For the future, we need to work more closely, to share innovations that bring more public confidence about privacy, yet deliver more research to improve lives. As a new non-ministerial department, ONS needs the support of the research community in showing Parliament the value of our shared aims.

Conclusion

There has never been a better time to innovate in research data access, and never has it been more important to do so.

Building a business case for SARs microdata from the 2011 Census

As part of our role in supporting the Samples of Anonymised Records from the 1991 and 2001 Censuses, it falls to the SARs support team at Manchester to assess demand for 2011 census microdata and to work towards fulfilling such needs. This process is underway, and you may have already provided us with information about how you have used SARs data and/or participated in our consultation survey, for which we thank you. Our next step is to put together a 'business case' for the SARs which we can present to the census offices, and later to funders. This is a vital step which requires a deeper understanding of how the SARs are used.

Keith Dugmore of Demographic Decisions will be working with us in this phase of the work. He will be seeking to undertake phone conversations with a number of users and research consumers. You may be contacted to give your views. Participation is voluntary, however we welcome information that helps us better meet data needs. Any queries can be directed to the SARs team.

Publications using the Samples of Anonymised Records

As part of the reporting requirements attached to the use of the SARs, we conduct an annual trawl of users' publications. The information we obtain is added to the SARs publication list, which is maintained by the Cathie Marsh Centre for Census and Survey Research at the University of Manchester and is available on our web site at www.ccsr.ac.uk/sars/publications/.

The publications list played a vital part in getting recognition for the importance of the SARs from the 2001 Census and this year will be especially important when we present our case for the SARs for 2011.

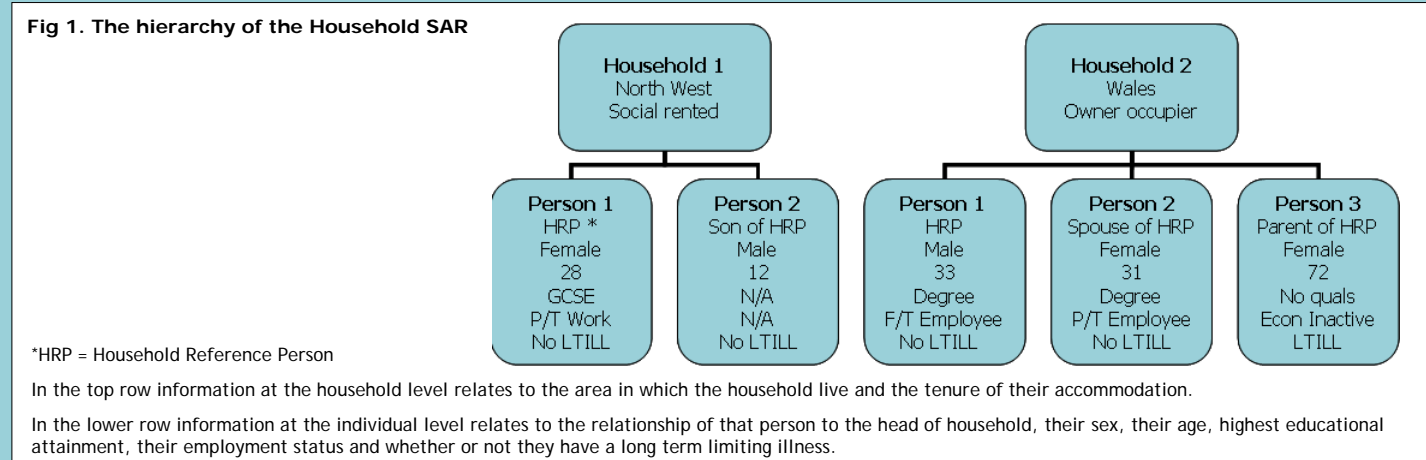
We would therefore be grateful if you could check your own publications on the publications list. Please email any amendments or updates to gillian.meadows@manchester.ac.uk, or provide the information on our web form at www.ccsr.ac.uk/sars/publications/

The Household Hierarchical Data

In both 1991 and 2001 Household SARs have been created. These files are samples of entire households, where household IDs have been preserved to enable analysts to link household members to each other.

A half-day workshop at the University of Newcastle in November aimed to demonstrate the power of the SARs for household analysis. The household files in particular are designed to enable analyses at both the individual and household level. With them, it is possible to:

- better describe households by drawing on information about the individual residents and summarising this for the entire household;
- better understand the family and household context that individuals live in, by using new household and family classifications to inform individual level analysis;
- better understand the impact of key relationships. It is, for example, possible to look at parent child relationships within the context of family units.



Working with hierarchical data requires an understanding of levels of analysis and knowledge of a few key data manipulation procedures within your choice of data analysis package. In SPSS, for example, the most important procedure is called aggregate; this allows you to produce a summary variable across cases with the same value of a break variable. In household analysis, this break variable is inevitably the household ID number.

Further guidance is available in a new step-by-step workbook for SPSS users, based on material used at a workshop at the University of Newcastle on 29 October (see www.ccsr.ac.uk/sars/events/). This workbook is available from the SARs web site and draws on a teaching dataset which is available for download under a standard End User Licence. More information can be found at www.ccsr.ac.uk/sars/support/.

More information about the household SAR including information on how to get access to the data under special licence is available at www.ccsr.ac.uk/sars/2001/hhold/.

Following last year's successful round of census data workshops, Census.ac.uk – the public face of the ESRC Census Programme – are running more. The first one of the season was held at Newcastle University at the end of October, and concentrated on household analysis using the Samples of Anonymised Records and was delivered by Jo Wathan from the Cathie Marsh Centre for Census and Survey Research.

Plans are in place for two workshops in early 2009 – both with a geographical focus – at the University of Southampton on 5 February and University of Essex on 9 March. The latter is being held as part of the ESRC Festival of Social Science 2009, which runs from 6-15 March.

Workshop booking and further information is at <http://census.ac.uk/workshops>.

The Census Programme supports a range of census datasets and services, including microdata, aggregate statistics, digitised boundaries, longitudinal data and interaction data, all of which will be covered by this series of workshops. To find out more about the data available via the Programme visit the web site at <http://census.ac.uk>

Also new this year is a dedicated mailing list, so potential workshop participants can be kept informed about forthcoming workshops and past attendees can give feedback on those they have already attended. To sign up, go to the Jiscmail web site at www.jiscmail.ac.uk/census-workshops

Forthcoming events

SARs user meeting, Royal Statistical Society, London, Thursday 23 April 2009. More details will be available in early 2009 from www.ccsr.ac.uk/sars/events/.

If you would like to host a workshop or ask us to give a seminar please contact us at sars-helpdesk@manchester.ac.uk or telephone 0161 275 4735.

CCSR short courses

CCSR run a programme of short courses on analysis methods relevant to users of the SARs. The complete list of short courses is available at www.ccsr.ac.uk/courses/.

Details of ESDS Government events can be found at www.ccsr.ac.uk/esds/events/.

Recent publications

Stillwell, J., S. Hussain and P. Norman (2008), The internal migration propensities and net migration patterns of ethnic groups in Britain, *Migration Letters*, 5, 2, 135-150.

Clarke, H. and S McKay (2008), Exploring disability, family formation and break up: reviewing the evidence, *DWP Research Report number 514*. The Department for Work and Pensions, London.

Bailey, N. and M. Livingston (2008), Selective migration and area deprivation: evidence from 2001 Census migration data for England and Scotland, *Urban Studies*, Vol. 45, No. 4, 943-961. Sage, London.

Finney, N. and L. Simpson (2008) Internal migration and ethnic groups: evidence for Britain from the 2001 Census, *Population, Space and Place*, 14: 63-83.



SARs team

Ed Fieldhouse	Director
Jo Wathan	Associate director: User support and development
Sam Smith	Associate director: IT
Gill Meadows	Administration
Angela Dale	Advisor

From January 2009 Selvino de Kort will be joining the SARs team in the role of Senior Development Officer.

SARs contacts

Email:
Sars-helpdesk@manchester.ac.uk

Telephone:
+44 (0)161 275 4735

Fax:
+44 (0)161 275 4722

SARs Support Team
CCSR
University of
Manchester
Humanities Bridgeford
Street
Manchester
M13 9PL

<http://census.ac.uk>