

Millennium Cohort Study:

Geographic Identifiers in MCS

June 2009

Jon Johnson

Centre for Longitudinal Studies

Faculty of Policy and Society

Institute of Education, University of London



Introduction

As Boyle et al¹ (2004) have noted “There has been a growing interest in the influence of local-area characteristics on the trajectories of individuals; a pertinent example is the ‘context versus composition’ health debate. However, the ability to investigate the influence of locale on individuals is hampered by the lack of geographically refined information that is usually included in longitudinal datasets”. They also recommended that where this data is made available below SAR-area it is “only be released in special downloads or safe settings”. The availability of these and further datasets in the future seeks to address this interest.

The UK's inconsistent geography has made it extremely challenging to produce and compare meaningful statistics over time. MCS is well placed to take advantage of the recent developments in geography led by ONS. The datasets being made available are designed to link into these ‘new’ geographies, Output Area, Census Area Statistic Ward.

Matching address at interview to areas longitudinally

Producing consistent geographical identifiers has hitherto been done via lookup tables of the postcode at interview to the required area. However, postcodes were never designed to be used as geo-spatial identifiers and as such are subject to change over time. For instance, matching in postcodes collected in MCS4 in 2008/9 with Census data collected in 2011 will produce a large number of postcodes that no longer match necessitating either going back to the original address or finding the postcode current at 2011 or using a translation table between the two.

To circumvent these problems a more straight-forward procedure has been adopted which allows us to produce consistent data more straight-forwardly using GIS.

Coding of Interview Address

Postcode at interview is checked against the National Statistics Postcode Directory² to ensure that the postcode is valid for the time period of the interview.

Six digit Eastings and 7 digit Northings (grid reference) are obtained from National Statistics Postcode Directory³ lookups closest to the interview date. This is equivalent

¹ Adding Geographical Variables and Identifiers to Longitudinal Datasets: A Report to the National Longitudinal Strategy Committee and ESRC 2004 (Paul Boyle, Andrew Cullis, Zhiqiang Feng, Robin Flowerdew, Vernon Gayle)

² National Statistics Postcode Directory (NSPD) The National Statistics Postcode Directory (NSPD) is maintained by the Office for National Statistics. Until May 2006 it was known as the All Fields Postcode Directory (AFPD).

³ National Statistics Postcode Directory (NSPD) The National Statistics Postcode Directory (NSPD) is maintained by the Office for National Statistics. Until May 2006 it was known as the All Fields Postcode Directory (AFPD).

to the address closest to the postcode mean, where postcodes are assigned to administrative areas using Ordnance Survey's Boundary-Line™ (a digitised set of administrative boundaries). England, Wales and Scotland grid-references are on a British National Grid basis.

Northern Ireland follows the same procedure but using the Central Postcode Directory obtained from NISRA. The Central Postcode Directory uses grid-references on the Irish National Grid basis.

This defines a fixed point to which digital maps of available boundaries can be overlaid to produce datasets based on which boundary the interview point resides within. Digital boundaries were obtained from UKBorders (<http://edina.ac.uk/ukborders/>). Appendix A outlines the data sources used and the geographies available at present.

Issues encountered

In areas where postcodes are more compact, the point will be by definition lay closer to the location of the interview than that in more dispersed areas, and users should be aware of this.

Discrepancies between lookups can occur for at least two reasons:

- a. the postcode boundary changes, and as a consequence the grid reference for the post code may change, this occurs when parts of postcodes are merged into others, compacting and expanding the area;
- b. errors are corrected / introduced between lookups

Availability of Datasets

Because of the potentially disclosive nature of these datasets, geographic identifiers below Government Office Region (GOR) are being released under Special Licence; these are listed at Appendix A. Please see the form available with the datasets for more details of the conditions of use.

Further Information

If you have any queries, please contact us at clsfeedback@ioe.ac.uk or view our website at <http://www.cls.ioe.ac.uk>.

APPENDIX A: Data Sources for Geographical Identifiers available under Special Licence

Dataset	Description	Boundary Data Source (UKBorders unless stated)
Interview Wards	1998 Ward Boundaries 2001 Census Area Statistic Ward ⁴	English Electoral Wards, 1998 Welsh Electoral Wards, 1998 Scottish Electoral Wards, 1998 Northern Ireland Electoral Wards, 2001 ⁵ English Census Area Statistic Wards, 2001 Welsh Census Area Statistic Wards, 2001 Scottish Census Area Statistic Wards, 2001 Northern Ireland Electoral Wards, 2001 ⁶
Interview Output Area (OA)	2001 Output Area ⁷	English Output Areas, 2001 Welsh Output Areas, 2001 Scottish Output Areas, 2001 Northern Ireland Output Areas, 2001

⁴ see http://www.statistics.gov.uk/geography/Statistical_CAS_ST_Wards.asp for definition and http://www.statistics.gov.uk/geography/geographic_area_listings/Statistical_CAS_wards.asp for code/name lookups

⁵ Northern Ireland Electoral Wards remained unchanged between 1998 and 2001

⁶ Northern Ireland does not have CAS wards but the electoral wards have alternative codes in Census outputs.

⁷ http://www.statistics.gov.uk/geography/census_geog.asp#oa

Dataset	Description	Boundary Data Source (UKBorders unless stated)
Interview Lower Super Output Area (LSOA)	2001 Output Area ⁸	English Output Areas, 2001 Welsh Output Areas, 2001 Scottish Output Areas, 2001 Northern Ireland Output Areas, 2001

⁸ http://www.statistics.gov.uk/geography/census_geog.asp#oa

APPENDIX B: Data Sources for Geographical Identifiers available with MCS Interview Data

Variables	Description	Data Source
Rural Urban	Rural Urban Classification England and Wales ⁹ , Scotland ¹⁰ , Northern Ireland ¹¹	English Census Area Statistic Wards, 2001 Welsh Census Area Statistic Wards, 2001 Scottish Census Datazones, 2001 Northern Ireland Electoral Wards, 2001
Country / Government Office Region	Country of Interview ¹² Government Office Region of Interview ¹³	English Government Offices for the Regions, 2001
Indices of Multiple Deprivation at 2001 Census	England Indices of Deprivation Welsh Index of Multiple Deprivation N. Ireland Multiple Deprivation Measure Scottish Index of Multiple Deprivation	ONS, England Indices of Deprivation (2004) Welsh Assembly, Welsh Index of Multiple Deprivation (2005) NISRA, Northern Ireland Multiple Deprivation Measure (2005) Scottish Govt, Scottish Index of Multiple Deprivation (2004)

⁹ ONS Rural and Urban Classification of CAS Wards 2005 see http://www.statistics.gov.uk/geography/downloads/Rural_Urban_Metadata.pdf

¹⁰ Scottish Executive Urban Rural Classification 2003-2004 see <http://www.scotland.gov.uk/Publications/2004/06/19498/38784>

¹¹ Statistical Classification and Delineation of Settlements 2005 see http://www.nisra.gov.uk/statistics/financeandpersonnel/DMB/urban_rural.html

¹² <http://www.statistics.gov.uk/geography/gor.asp>

¹³ <http://www.statistics.gov.uk/geography/gor.asp>

Version History

June 2009