

Millennium Cohort Study:

**Geographic Identifiers in MCS**  
(2<sup>nd</sup> Edition)

December 2010

Jon Johnson

**Centre for Longitudinal Studies**

Institute of Education, University of London



## **Introduction**

As Boyle et al<sup>1</sup> (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 Directory<sup>2</sup> 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 Directory<sup>3</sup> lookups closest to the interview date. This is equivalent

---

<sup>1</sup> 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)

<sup>2</sup> 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).

<sup>3</sup> 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.

Appendix B lists the contents of the Special Licence Datasets.

Appendix C lists the geographies available on the Datasets available under the standard End User Licence.

### **Further Information**

If you have any queries, please contact us at [clsfeedback@ioe.ac.uk](mailto: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 <sup>4</sup>	English Electoral Wards, 1998 Welsh Electoral Wards, 1998 Scottish Electoral Wards, 1998 Northern Ireland Electoral Wards, 2001 <sup>5</sup>  English Census Area Statistic Wards, 2001 Welsh Census Area Statistic Wards, 2001 Scottish Census Area Statistic Wards, 2001 Northern Ireland Electoral Wards, 2001 <sup>6</sup>
Interview Output Area (OA)	2001 Output Area <sup>7</sup>	English Output Areas, 2001 Welsh Output Areas, 2001 Scottish Output Areas, 2001 Northern Ireland Output Areas, 2001

<sup>4</sup> see [http://www.statistics.gov.uk/geography/Statistical\\_CAS\\_ST\\_Wards.asp](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](http://www.statistics.gov.uk/geography/geographic_area_listings/Statistical_CAS_wards.asp) for code/name lookups

<sup>5</sup> Northern Ireland Electoral Wards remained unchanged between 1998 and 2001

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

<sup>7</sup> [http://www.statistics.gov.uk/geography/census\\_geog.asp#oa](http://www.statistics.gov.uk/geography/census_geog.asp#oa)

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

---

<sup>8</sup> [http://www.statistics.gov.uk/geography/census\\_geog.asp#oa](http://www.statistics.gov.uk/geography/census_geog.asp#oa)

## APPENDIX B: Variables available under Special Licence

### Geographical Identifiers, Ward Level (SN 5728)

Variable name	Description
aw98nam	S1 Ward (1998 Boundaries)
award98	S1 Ward Code (1998 Boundaries)
acasward	S1 Census Statistic Ward Code
acwardnm	S1 Census Area Statistic Ward
bw98nam	S2 Ward (1998 Boundaries)
bward98	S2 Ward Code (1998 Boundaries)
bcasward	S2 Census Statistic Ward Code
bcwardnm	S2 Census Area Statistic Ward
cw98nam	S3 Ward (1998 Boundaries)
cward98	S3 Ward Code (1998 Boundaries)
ccasward	S3 Census Statistic Ward Code
ccwardnm	S3 Census Area Statistic Ward
dw98nam	S4 Ward (1998 Boundaries)
dward98	S4 Ward Code (1998 Boundaries)
dcasward	S4 Census Statistic Ward Code
dcwardnm	S4 Census Area Statistic Ward

### Geographical Identifiers, Lower Super Output Area Level (SN 5729)

Variable name	Description
alsoa	S1 Lower Super Output Area Code
blsoa	S2 Lower Super Output Area Code
clsoa	S3 Lower Super Output Area Code
dlsoa	S4 Lower Super Output Area Code

### Geographical Identifiers, Output Area Level (SN 5730)

Variable name	Description
aoa	S1 Output Area
boa	S2 Output Area
coa	S3 Output Area
doa	S4 Output Area

**APPENDIX C: Data Sources for Geographical Identifiers available with MCS Interview Data**

<b>Variables</b>	<b>Description</b>	<b>Data Source</b>
Rural Urban	Rural Urban Classification England and Wales <sup>9</sup> , Scotland <sup>10</sup> , Northern Ireland <sup>11</sup>	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 <sup>12</sup> Government Office Region of Interview <sup>13</sup>	English Government Offices for the Regions, 2001

<sup>9</sup> ONS Rural and Urban Classification of CAS Wards 2005 see [http://www.statistics.gov.uk/geography/downloads/Rural\\_Urban\\_Metadata.pdf](http://www.statistics.gov.uk/geography/downloads/Rural_Urban_Metadata.pdf)

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

<sup>11</sup> Statistical Classification and Delineation of Settlements 2005 see [http://www.nisra.gov.uk/statistics/financeandpersonnel/DMB/urban\\_rural.html](http://www.nisra.gov.uk/statistics/financeandpersonnel/DMB/urban_rural.html)

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

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