



Geography and Longitudinal Data

Understanding Society
The UK Household Longitudinal Study

Dr Gundi Knies

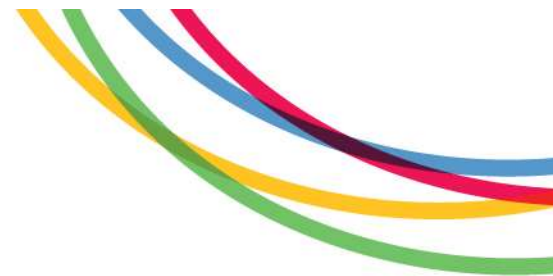
ISER University of Essex
gknies@essex.ac.uk



An initiative by the Economic and Social Research Council, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research.

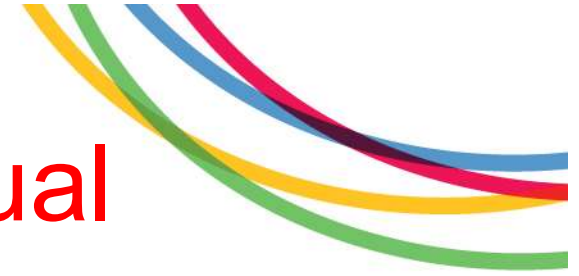


What is *Understanding Society*?



- High quality longitudinal data aimed to help us understand the short, medium and long term effects of social and economic change
- The Study collects both *objective* and *subjective* indicators of different aspects of life
- It offers opportunities for research across multiple disciplines (e.g. sociology, economics, geography, psychology and health sciences)

Topics covered in the annual interviews...



Key Topics

- Education
- Employment
- Family and household
- Health, health behaviours, wellbeing
- Income, housing, wealth, expenditure & deprivation
- Ethnicity

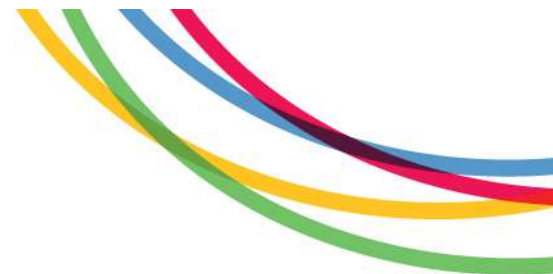
Supporting topics

- Neighbourhood characteristics
- Preferences, expectations across key topic domains
- Social networks, support, reciprocity
- Transport
- Time Use
- Personality traits, identity, beliefs

Additional topics

- Environmental behaviours
- Political behaviour
- Leisure activities

Household panel design

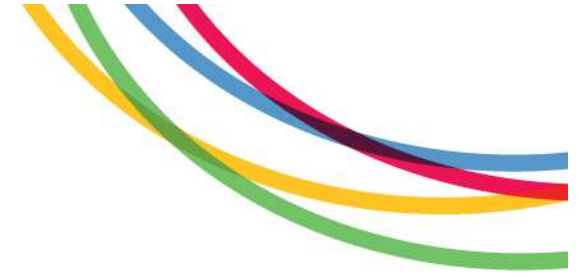


UKHLS is a longitudinal sample of individuals representing the whole UK population, and interviewed within a household context:

- starts with randomly selected addresses
- collects information about all residents at selected addresses
- follows the sample members' life courses over time, collecting data also from people living with them as they move and form new households



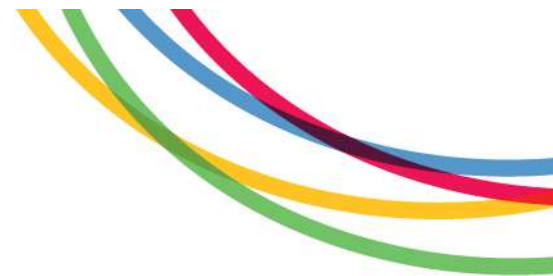
Target population and samples (1)



The British Household Panel Survey (BHPS) element of Understanding Society:

- 1991 - began with a representative probability sample of the residential population living in private households in Britain
- 1997 - Scotland and Wales boost samples added
- 2001 - Northern Ireland sample added -> BHPS became representative of the residential population living in private households in the United Kingdom.

Target population and samples (2)



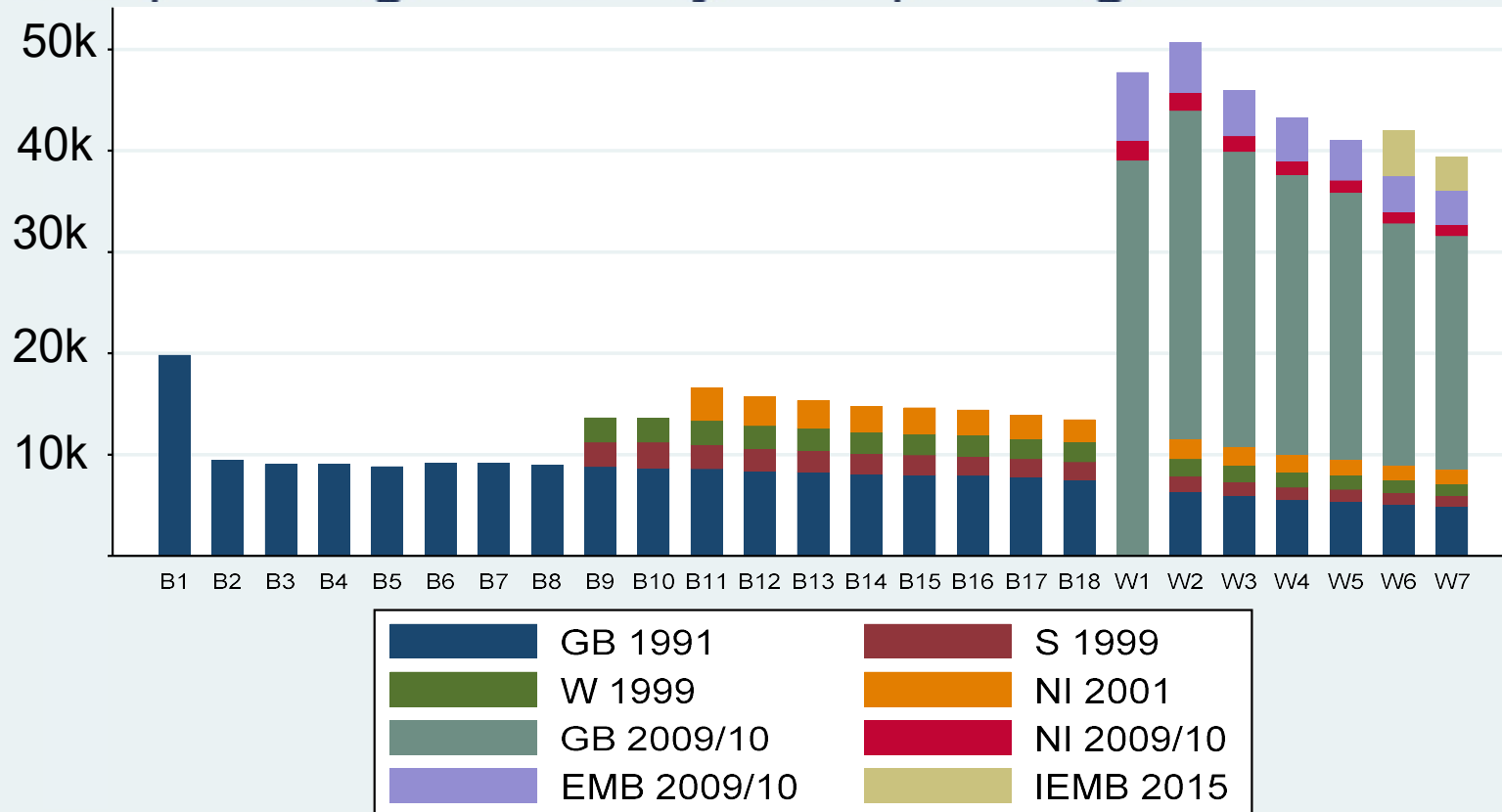
The new household longitudinal study element of Understanding Society:

- 2009/10: Very large (26k households!), regionally representative sample of the UK's General Population (GP)
- 2009/10: PLUS Ethnic Minority* Boost (EMB) sample
- 2015: EMB refresher, additionally boosting immigrants to the UK (IEMB - Immigrant and Ethnic Minority sample)

* Target groups: Indian, Pakistani, Bangladeshi, Caribbean, and African

Responding adults over time

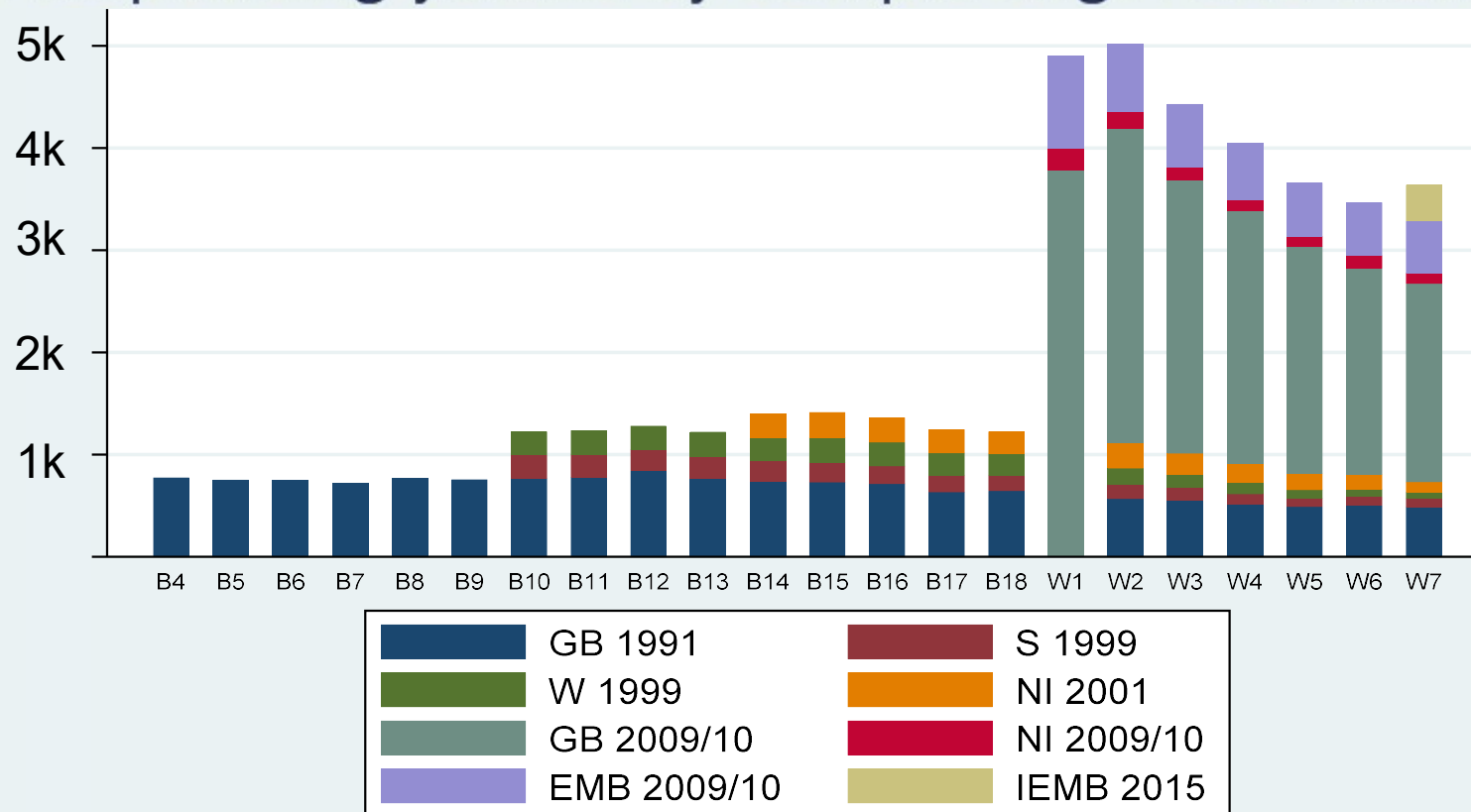
Responding adults by sample origin and wave



Source: Understanding Society, 2018, 7.1

Responding youths over time

Responding youths by sample origin and wave



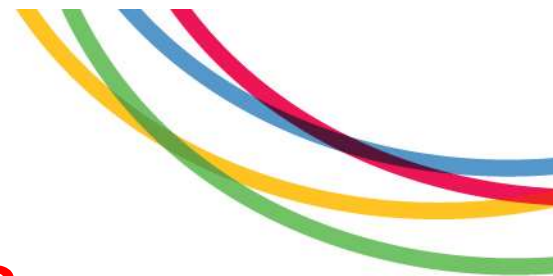
Source: Understanding Society, 2018, 7.1

Data collection



- Most of the data is collected using **face-to-face** interviews with adults (aged 16+) and **self-completion** interviews with children aged 10-15
- Questions are repeated, but not every question appears in every wave
- More than just survey questions: Biological specimens and data linkage
such as to UK Census

From addresses to geographies and indicators



- Start with sample households' exact addresses at the time of interview/fieldwork
- Match postcodes with the May release of that year's ONS Postcode Directory containing:
 - Administrative, electoral, and health unit identifiers, and population-weighted centroids
 - Official neighbourhood classifications
 - For full list see ONSPD user guidance <http://geoportal.statistics.gov.uk/>

Geographical data available (1)

Already included IDs/characteristics			
core	6614	EUL	Understanding Society: GOR_DV & URBAN_DV
core	6931	SL	Understanding Society: GOR_DV & URBAN_DV
linked	7454	SL	Census 2001 Rural-Urban Indicators
linked	7630	SL	Census 2011 Rural-Urban Indicators
linked	6674	SL	Census 2001 Output Area Classification
linked	7629	SL	Census 2011 Output Area Classification
linked	7453	SL	Acorn Types
linked	7533	SL	Geographical Accessibility (Waves 1-3)

Replace \$\$\$\$ by the study number to get to the data basket:

[https://discover.ukdataservice.ac.uk/catalogue/?sn=\\$\\$\\$\\$](https://discover.ukdataservice.ac.uk/catalogue/?sn=$$$$)

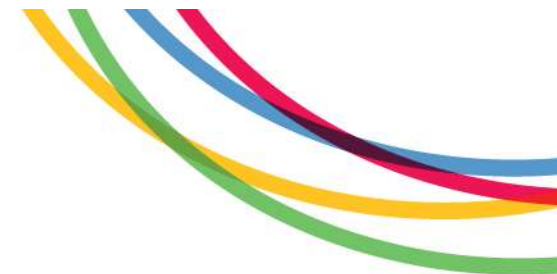
Geographical data available (2)

Merge your own geo-coded data using...

core	6676	SEC	Understanding Society: postcode/Grid ref
link-id	7182	SL	School Codes (Wave 1)
link-id	6666	SL	Local Authority District
link-id	6668	SL	Westminster Parliamentary Constituencies
link-id	6671	SL	Local Education Authorities
link-id	6675	SL	Travel to Work Areas
link-id	6672	SL	Strategic Health Authorities
link-id	6673	SL	Primary Care Organisations
link-id	6669	SL	Census 2001 Area Statistics Wards
link-id	7245	SL	Census 2001 Middle Layer Super Output Areas
link-id	7249	SL	Census 2011 Middle Layer Super Output Areas
link-id	6670	SL	Census 2001 Lower Layer Super Output Areas
link-id	7248	SL	Census 2011 Lower Layer Super Output Areas

Test-run your external data merge on ONSPD

Produce results for small geographies



Geography	Number in UK	Number in sample	Sample members in unit (median and IQR)
OA 2001	223,019	26,446	2 (1;3)
LSOA 2001	34,379	13,437	4 (2;6)
MSOA 01	7,195	4,783	11 (6;17)
TTWA	243	228	145 (64;321)
LAD	219	217	201 (112;330)
GOR	12	12	5,443 (3,829;6,784)

Base: Wave 1 Responding adults and children in their household

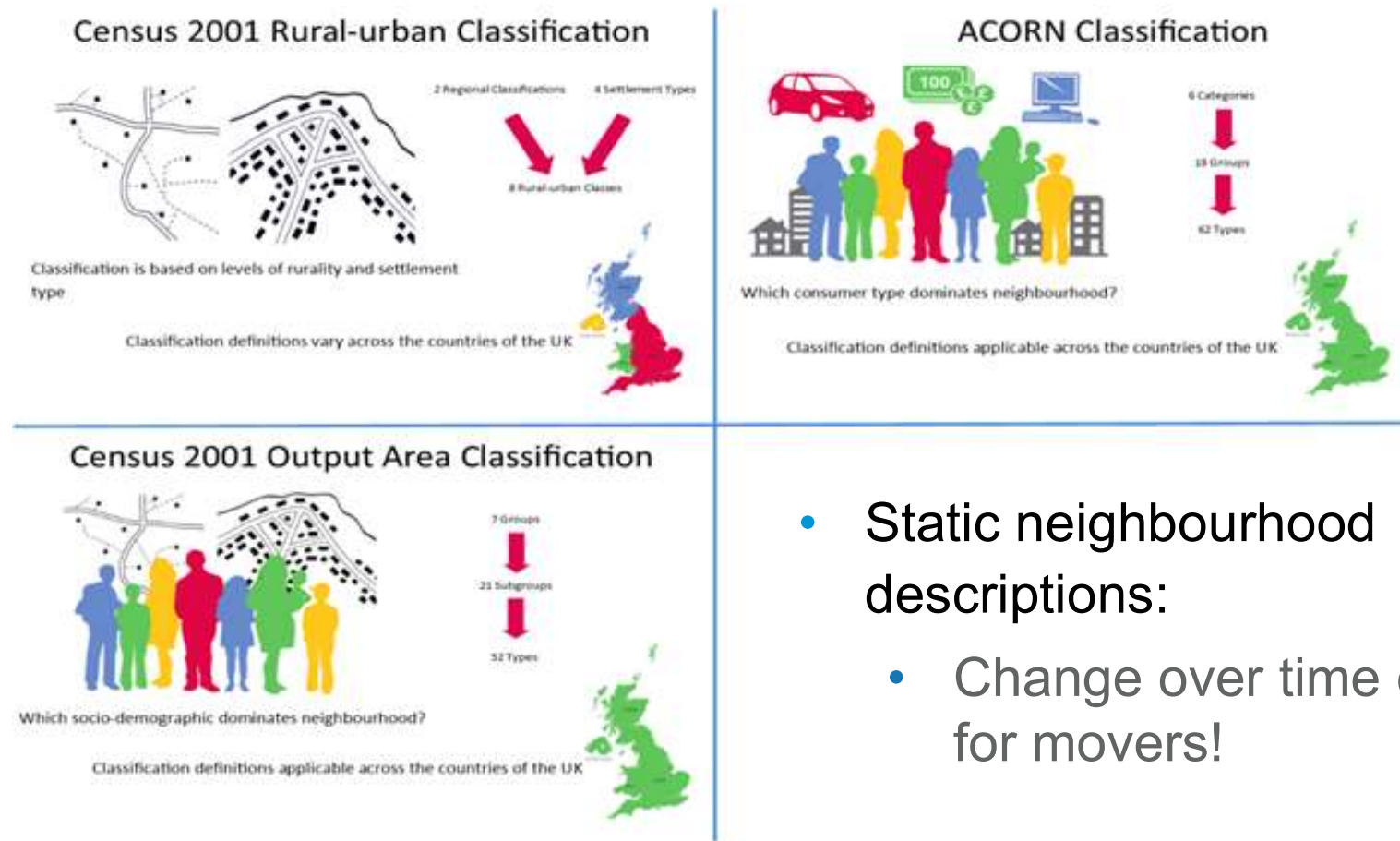
A lot of longitudinal research exploits change over time ...



Change over time in the neighbourhood context can occur because:

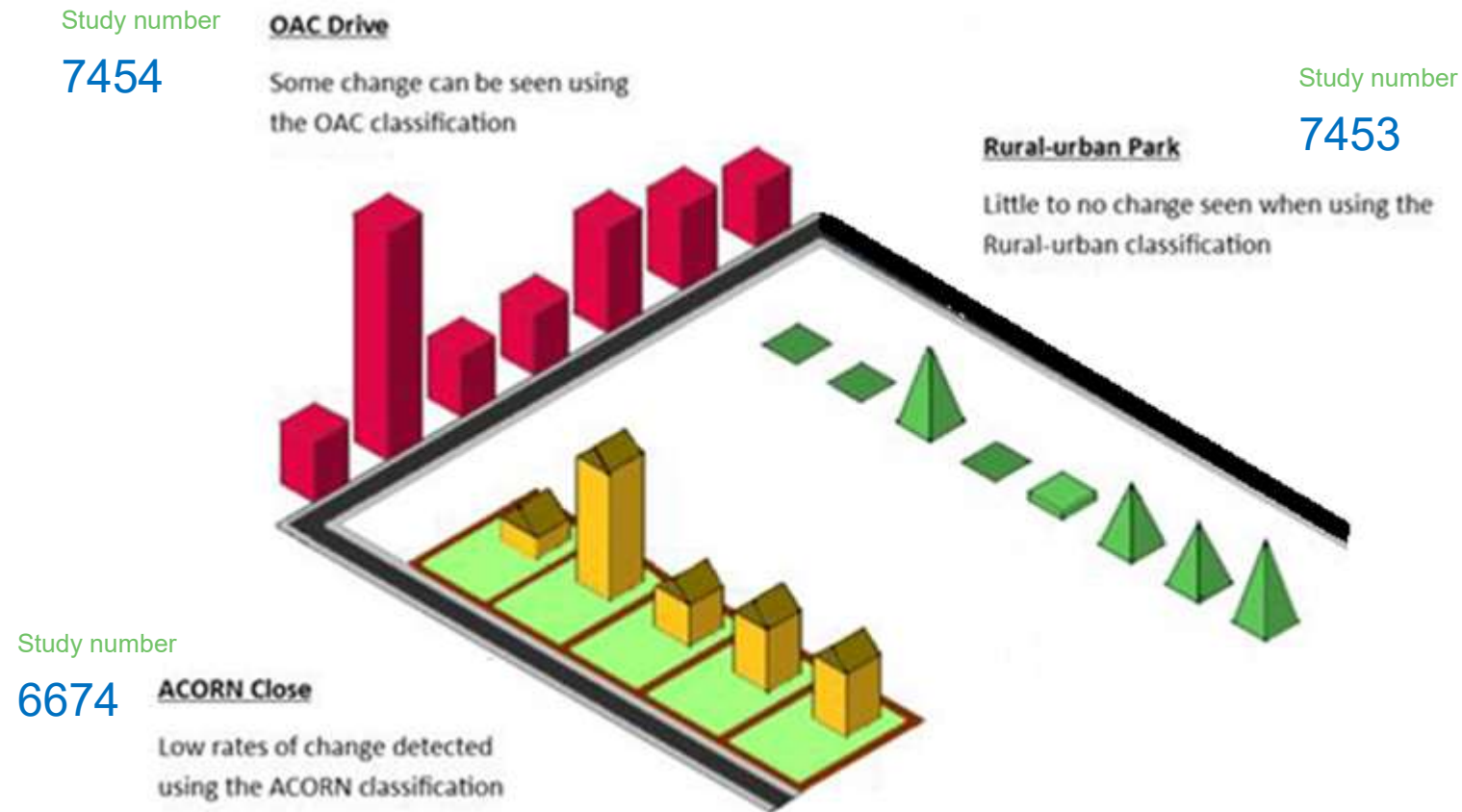
- Individual moved to a different place (may be in the same micro-area though!)
- Neighbourhood context genuinely changed
- The neighbourhood boundary changed over time
- The neighbourhood classification has changed

Rates of change in neighbourhood conditions



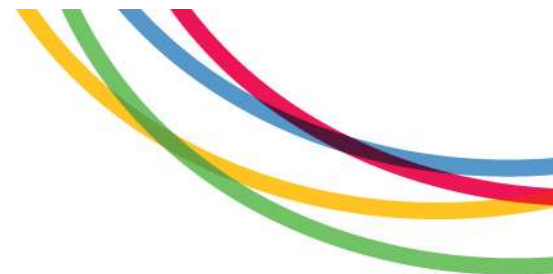
- Static neighbourhood descriptions:
 - Change over time only for movers!

Rates of change in neighbourhood conditions (Wave 1 to Wave 2)



Source: Knies (2017): Exploring the value of Understanding Society for Neighbourhood effects Research.
DOI: 10.1163/24523666-01000006

Weblinks for further info



Understanding Society Website

<https://www.understandingsociety.ac.uk/>

Data access routes

<https://www.understandingsociety.ac.uk/documentation/access-data>

Online Documentation

<https://www.understandingsociety.ac.uk/documentation>

Understanding Society on UKDS website

<https://discover.ukdataservice.ac.uk/series/?sn=2000053>

Geography in the ONS Longitudinal Study

UK Data Service



Dr Oliver Duke-Williams

Department of Information Studies, UCL



Webinar July 18th 2018: Geography and longitudinal data - the UK Household Longitudinal Study and the UK Longitudinal Studies

Contact: o.duke-williams@ucl.ac.uk, [@oliverdw](https://twitter.com/oliverdw); celsius@ucl.ac.uk, [@celsiusnews](https://twitter.com/celsiusnews)

About the ONS LS

- The ONS LS is based on 4 sample birth dates
 - i.e. 1% of population
 - Sample members do not know that they are in the sample
- The data for England and Wales include both census data (1971-2011) and administrative data
- Access to the data is controlled
- [Further info](#) if needed in following hidden slides

Longitudinal data

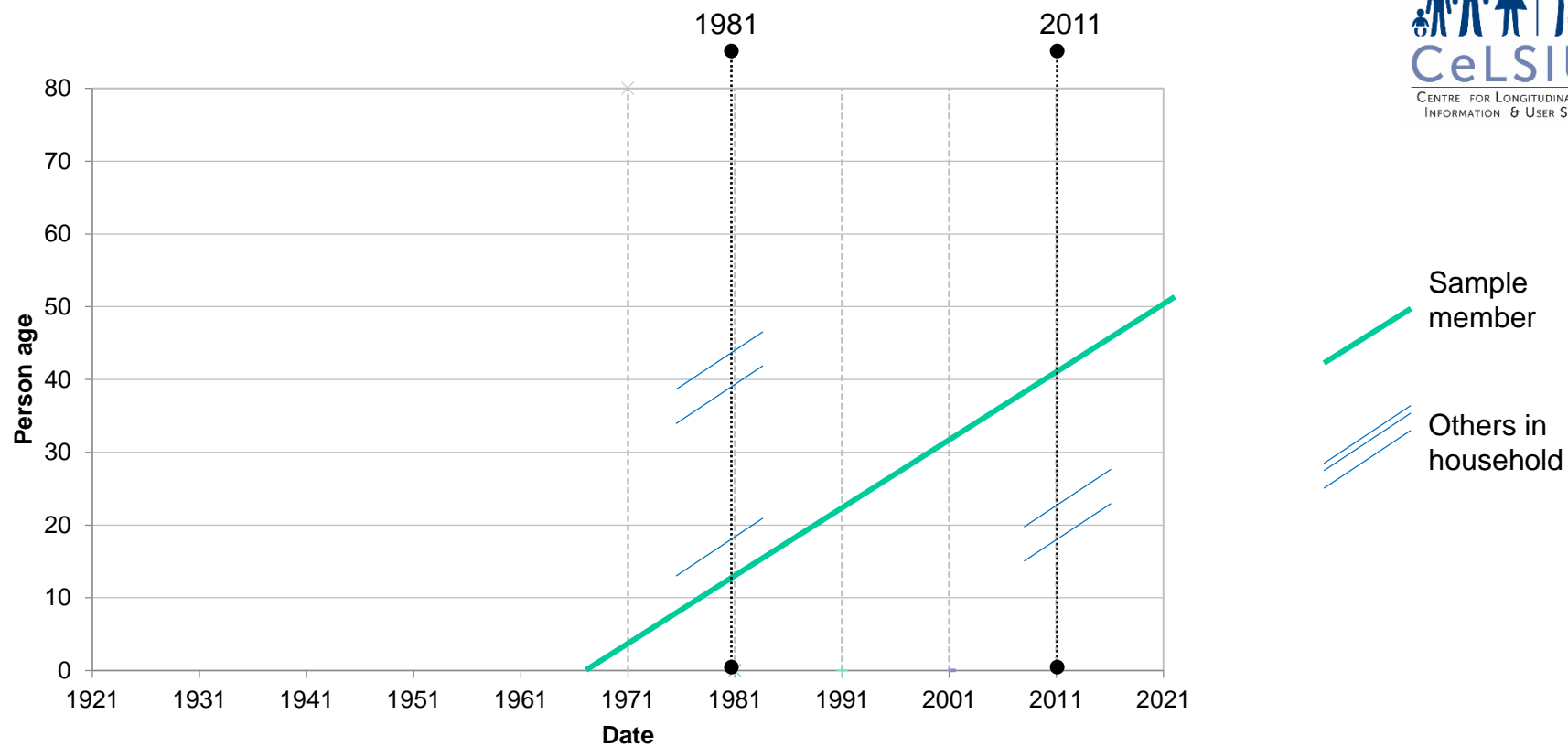
- There are three longitudinal studies in the UK
- They have different sample sizes and cover different time periods
- They also differ in the range and amount of linked data
- All have secure access arrangements

UK Census Longitudinal Studies

	Sample size	1971	1981	1991	2001	2011
ONS Longitudinal Study	4/365.25	✓	✓	✓	✓	✓
Scottish Longitudinal Study	20/365.25		✓	✓	✓	✓
Northern Ireland Longitudinal Study	104/365.25			✓	✓	✓

What is in the LS?

- Similarly to census microdata, all variables apart from individual identifiers
 - More detail than the safeguarded and open microdata files
- Sample members and others in the household



Census Data



From each census

LS sample members

Other household members

Age, sex, marital status, country of birth

Family and household types, communal establishments

Housing: tenure, rooms and amenities

Qualifications, economic activity, occupation, industry and social class

Travel to work and one-year migration

Geographical information

2011 census

Ethnicity (1991-2011)

National identity (2011)

Year of arrival (2011)

Limiting long-term illness (1991-2011) & self-rated health (2001, 2011)

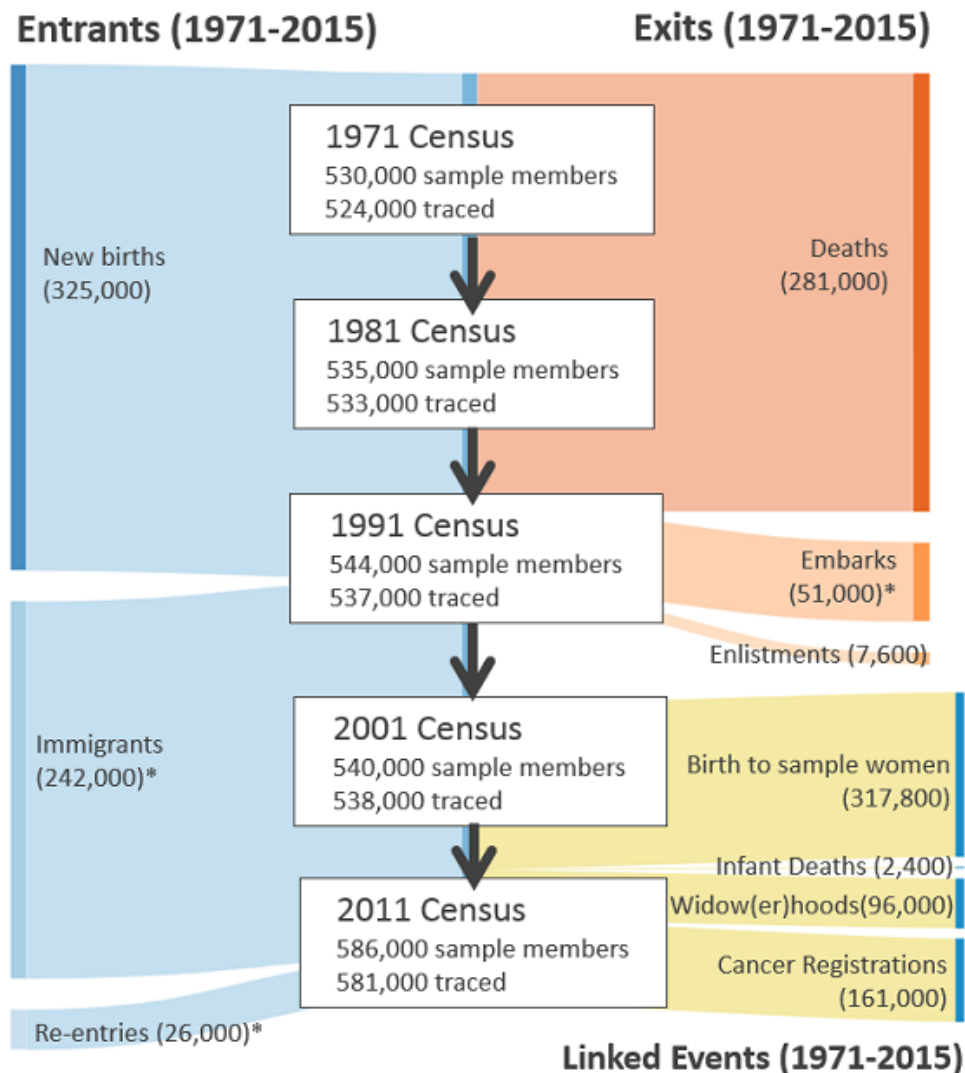
Care-giving (2001, 2011)

Religion (2001, 2011)

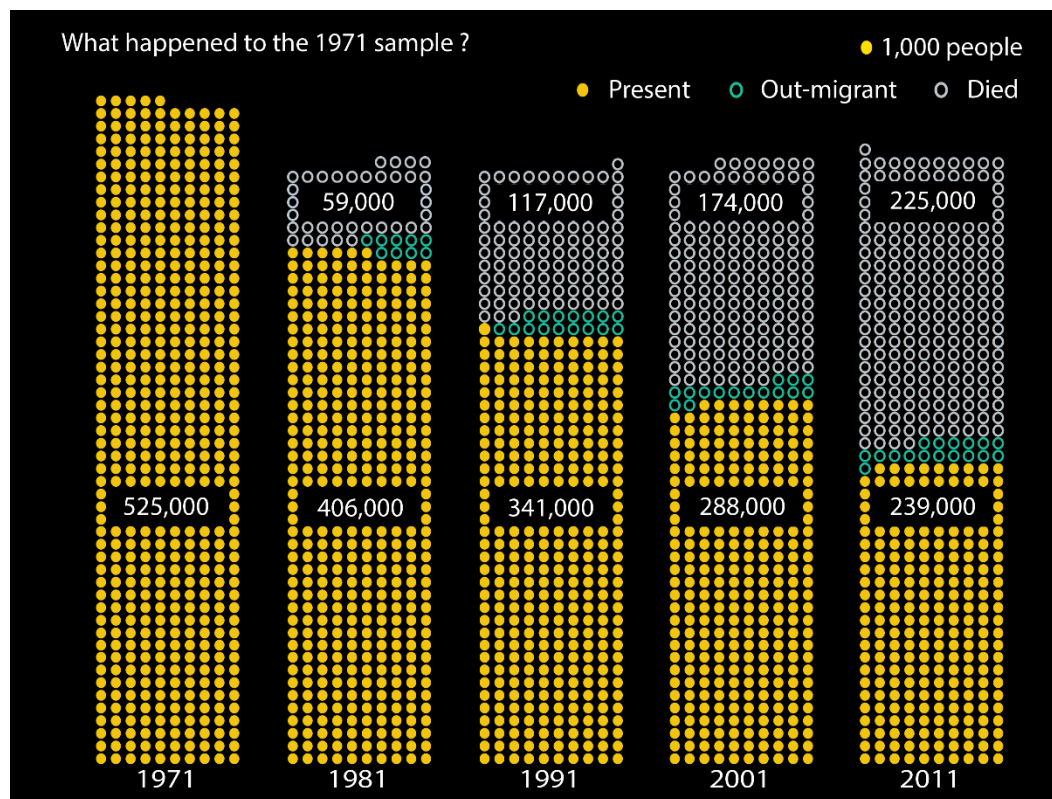
Short-term migration (2011)

Main language

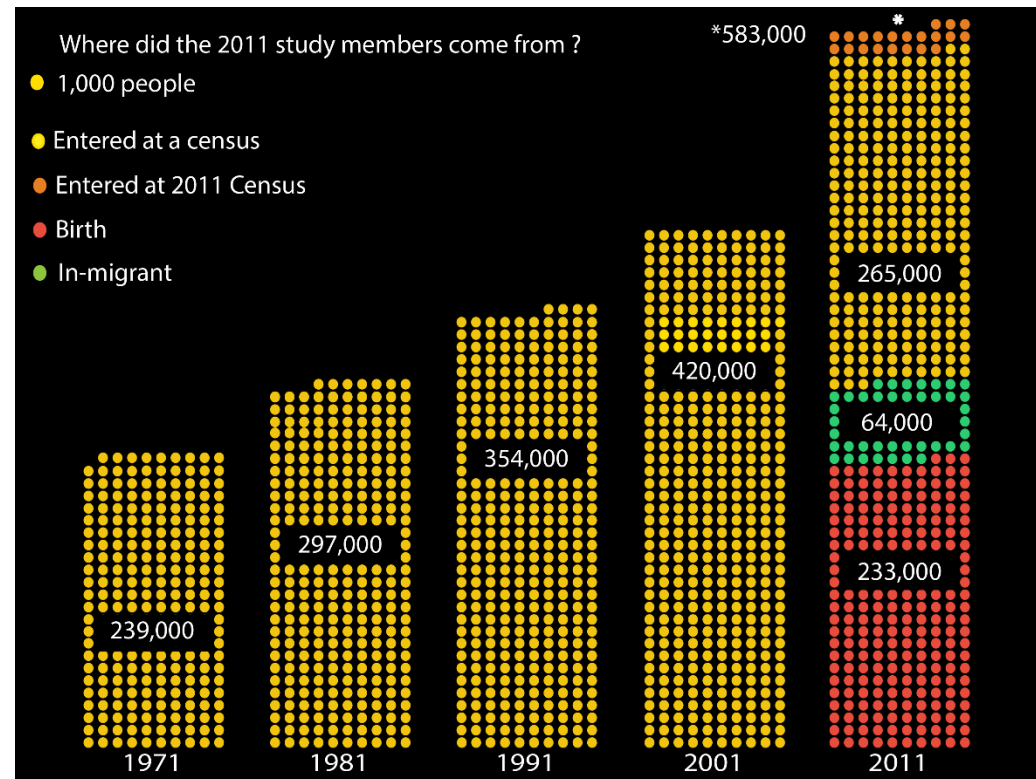
LS structure: England & Wales



Following on from 1971...



...leading up to 2011



Developments since 2012

Beta test projects / 2011 linked data launch

User survey - matching users

Synthetic data

ONS consultations: 2021 Census, new approved researcher scheme & new VML

Training and Roadshows

Welsh linked data launch

Data dictionaries LS and UKLS

Geography in the LS

- The LS consists of multiple files, most of which contain a variety of geographic identifiers
- We will concentrate on the census files
- The most detailed geographies are contained within restricted access 'x files' tables
 - Researchers can not use these for standard analysis or for reporting results
 - They can be used for linking, etc

Which geography?

There are a variety of types of geography, including

- Place of enumeration
- Place of usual residence
- Place of second residence
- Place of work
- Place of usual residence in the past
- Student's term-time address
- Country of birth
- Place of birth / place of enumeration in 1939

Poll: country of birth



To the nearest thousand, how many people (in the 2011 Census, England and Wales) were born in Croatia?

(The LS will contain approx. 1% of these people)

- 12,000
- 10,000
- 8,000
- 6,000

1971

(pre~ and post~ 1974 variables)

- Standard Region, County, District
- Local Authority
- New Town
- Regional Health Authority, Area Health Authority, Health District
- (Ward)
- (Grid reference)

1981

- Standard Region, County, District
- Local Authority
- New Town
- Regional Health Authority, Area Health Authority, Health District
- Travel To Work Area
- (Ward, Enumeration District)

1991

- Standard Region, County, District
- Local Authority
- New Town
- Regional Health Authority, Area Health Authority, Health District
- Travel To Work Area
- (Ward, Enumeration District)
- (Postcode)

2001

- Government Office Region, County, District
- Parliamentary constituency, European constituency
- National Park
- Regional Health Authority, Area Health Authority, Health District
- Travel To Work Area
- (Ward, Output Area)
- (Parish)
- (Primary Care Group, Primary Care Trust)
- (Grid reference)

2011

- Government Office Region, County, District
- National Park
- TTWA
- 1991 District codes

- (Ward, Output Area, LSOA, MSOA)
- (Postcode)
- (Second Residence OA, Postcode etc)
- (Workplace Zone)
- (Grid reference)

Using low-level geographies (1)

It is relatively easy for data to be re-coded on the basis of an underlying unit (OA, LSOA, ward etc)

Using low level geographies (2)

It is possible to use OA or ward geographies to attach contextual (area-level) data

- User can submit a data set with all OAs etc, and matching values will be selected
- It must not be possible to deduce a sub-district location from the final results
- Usual approach is to convert small area values from raw values to deciles etc

Using the LS

The LS is free at the point of use

- Use in person at the ONS secure setting
- Use by remotely submitting Stata, SPSS etc code
- More info: ucl.ac.uk/celsius and calls.ac.uk

Application process

- Needs a combination of Accredited Researcher and approved project
- Apply via ucl.ac.uk/celsius who will assist and advise on form completion etc

