

---

# Secondary analysis workshop: Day 2: Starting secondary analysis – linking data files

Sarah King-Hele & Hersh Mann  
UK Data Service

7 March 2014

---

UK Data Service

---



---

# Overview

- A quick introduction to types of survey structures and why may need to link files
- Practical exercises with EHS and UKHLS data



---

## Key data: [UK surveys](#)

- Labour Force Survey
- Annual Population Survey
- Health Surveys (England/Wales/Scotland)
- General Household Survey/General Lifestyle Survey
- Expenditure and Food Survey/Living Costs and Food Survey
- Crime Survey for England and Wales
- Family Resources Survey
- ONS Omnibus/Opinions Survey
- Citizenship Survey
- British Social Attitudes Survey



---

# UK Surveys

- data about individuals or households
- often commissioned by the Government and conducted by the ONS or NatCen
- large sample sizes
- nationally-representative
- repeated cross-sectional surveys – a new sample of people every time the survey is run
- many surveys are repeated every year
- Well documented
- cover many topics including health, work, crime, social attitudes, family expenditure, living costs, housing etc.



---

# Longitudinal data

Similar to 'UK Surveys'

- individual level data
- large samples, nationally-representative
- surveys are repeated, often annually

BUT

- they follow the same individuals over time
- new respondents are added regularly to keep numbers up



---

## Key data: Longitudinal / Panel studies

- National Child Development Study (NCDS)
- 1970 British Cohort Study (BCS70)
- Millennium Cohort Study (MCS)
- British Household Panel Survey (BHPS)
- Understanding Society (UKHLS)
- English Longitudinal Study of Ageing (ELSA)
- Growing Up in Scotland (GUS)
- Longitudinal Study of Young People in England (LSYPE)



---

# British Household Panel Survey

- Collected and deposited by the Institute for Social and Economic Research at the University of Essex
- Follows the members of 5500 households first sampled in 1991 every year to 2008
- Coverage includes:  
income, labour market behaviour, social and political values, health, education, housing and household organisation



---

# Understanding Society

- The follow-up to the BHPS. Expanded to 40,000 households and 100,000 individuals including the original BHPS sample
- coverage includes:
  - current employment and earnings; employment status; parenting and childcare arrangements; family networks; benefit payments; political party identification; household finances; environmental behaviours; consents to administrative data linkage (health and education).
- incorporates an innovation panel
- collects anthropometric data



---

# Linking data files

- Cohort and panel studies data are often organised across a number of different data files within waves as well as across waves.
- Questions are not always repeated for each wave (e.g. birth date does not change) but researchers still want to use that information in subsequent sweeps. To do this the data files must be linked using a common variable.
- How is this done? We will now learn to do this with EHS and UKHLS data

