The University Of Sheffield.

The Longitudinal Study of Young People in England

Daniel Gladwell (ScHARR)
Steven McIntosh (Economics)

CWIPP seminar: 21st May 2014

Overview

The dataset

- LSYPE: longitudinal, young people, England.

- Began in 2004, aged 13/14.

- Follow up surveys annually for 7 years in total until 2010.

- Current proposal to do an age 25 survey later this year.

Years, waves & ages

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
Calendar Year	2004	2005	2006	2007	2008	2009	2010
School Year	Yr 9	Yr 10	Yr 11	Yr 12	Yr 13		
Cohort Member Age	13/14	14/15	15/16	16/17	17/18	18/19	19/20
Key Stage	KS3		GCSE		A level		

Sampling

- PSU are schools (all schools included)
- Deprived schools over-sampled
- Pupils sampled within schools
- Ethnic minority pupils over-sampled
- Modal group is 21 sampled pupils in a school
- Maximum sample size within schools is 41

Category	Proportion
Ethnicity	
White	83.0%
Bangladesh	1.0%
Pakistan	2.3%
Indian	2.3%
African	1.8%
Caribbean	1.4%
Mixed	2.2%
Other	2.5%
Not obtained	3.5%
GOR	
North East	5.3%
North West	14.8%
Yorkshire and The Humber	10.6%
East Midlands	8.9%
West Midlands	11.4%
East of England	11.0%
London	12.7%
South East	15.5%
South West	9.8%
Qualifications	
Not achieved level 1	10.5%
Achieved level 1, but not 2	33.0%
Achieved level 2	56.5%
Sex	
Male	51.0%
Female	49.0%

Attrition

- 15,770 young people originally in data set.

- Nationally representative.

- Attrition rates very low: response rates of 74%, 86%, 92%, 92%, 89%, 87% and 90%.

- Sample size in wave 7:   8,700 (54% of baseline)

- Weights to keep sample representative

Survey Methods

- Waves 1-4 face-to-face. Waves 5-7 use multi-mode approach

- Waves 1-4 include parents' questionnaire

- No school questionnaire

Young person questionnaire

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
1. Employment & Resources							
Current activities				•	•	•	•
Activity history				•	•	•	•
Jobs and training				•	•	•	•
Apprenticeships		•	•	•	•	•	•
NEET				•	•	•	•
Parental employment			•				
Income and benefits				•	•	•	•
Education Maintenance Allowance (EMA)			•	•	•		
Job search					•	•	•
Attitudes to work						•	•
Attitudes to debt				•	•	•	

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
2. Household Formation & Relationships							
Relations with parents	●						
Family stability	●	●	●				
Relationships and sexuality						●	●
Childcare and caring responsibilities	●	●	●	●	●	●	●
Care to learn					●	●	
Neighbourhood context					●		
3. Education							
Qualifications being studied for				●	●	●	●
Higher education plans	●	●	●	●	●	●	●
Higher education					●	●	●
Attitudes to higher education						●	●
Attitudes to school/teachers	●	●	●	●			
Rules and discipline	●						

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
Academic self-concept	•	•	•				
Subject choices and reasons	•	•					
Homework	•	•					
ICT	•	•					
Study support & Extra classes	•	•	•				
Future plans and advice	•	•	•	•			
Information, advice and guidance	•	•	•	•		•	•
4. Health & Wellbeing							
Health and disability		•	•	•		•	•
Mental and emotional health		•		•			•
Health behaviours / Risk taking (smoking, drinking, drugs)	•	•	•	•		•	•
5. Identity & Participation							
Antisocial risk taking behaviours / Crime (truancy, exclusions, bullying & discrimination, antisocial activities, contact with police)	•	•	•	•	•	•	•
Social Attitudes					•		

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
Locus of Control		•		•			•
Participation	•	•	•	•	•		•
Sport frequency	•	•	•	•		•	•
Use of leisure time	•	•		•			
Other							
Demographics	•	•	•	•	•	•	•
SIC/SOC codes				•	•	•	•
Data linkage consent				•	•	•	•
Future contact details			•	•	•	•	•

Parental Questionnaires

- Education, employment earnings etc
- Information about child (SEN, school history, extra-curricular activities etc)
- Expectations and aspirations for child.

Variable Names

- Consistent naming style
- [Prefix 1] [Question name] [Suffix 1] {[Suffix 2]}
- Prefix 1 is the wave (w1, w2 etc)
- Suffix 1 is the respondent (yp, mp, sp)
- Suffix 2 is to identify multi-coded responses.
- E.g. w3sencurrmp, w2healcyp

Merging in Data from Other Sources

- Currently 3 admin data sets can be merged into LSYPE, within SDS

- National Pupil Database – KS2, 3 and 4 attainment, SEN, FSM

- Schools Database – type of school, size, PTR, school level attainment, SEN, language

- Geographical data – National Statistics Postcode Directory - local deprivation indices.

Strengths of LSYPE

- Longitudinal
- Follow young people from education into labour market
- Annual observations (unlike birth cohorts)
- Detailed attainment data
- Family background and parental information
- Attitudinal data
- Merge in admin data

Weaknesses of LSYPE

- Lack of data pre-14 (except KS2)
- No teacher reported data
- No parental data after wave 4
- Health, well-being and non-cognitive questions not asked every year.
- Just a single cohort

LSYPE 2

- A second LSYPE cohort were started in 2013.

- Will be the GCSE cohort in summer 2015

- No data (or details) available yet

Accessing the Data

Gaining an Overview (1)

- Key documentation can be accessed at the UK data archive

http://discover.ukdataservice.ac.uk/catalogue/?sn=5545

DOCUMENTATION

Title	File Name	Size (KB)
KS2 KS3 KS4 Data Dictionary Spreadsheet	5545_ks2_ks3_ks4_data_dictionary.xlsx	27
KS5 Data Dictionary Spreadsheet	5545_ks5_data_dictionary.xlsx	69
LSYPE User Guide Waves One to Seven	5545lsype_user_guide_wave_1_to_wave_7.pdf	1460
Unclassified	5545removing_data_where_respondents_dropped_out.pdf	250
Wave Five documentation	5545wave_five_documentation.pdf	2547
Wave Four documentation	5545wave_four_documentation.pdf	2021
Wave One documentation	5545wave_one_documentation.pdf	1287
Wave Seven documentation	5545wave_seven_documentation.pdf	2279
Wave Six documentation	5545wave_six_documentation.pdf	1886
Wave Three documentation	5545wave_three_documentation.pdf	2068
Wave Two documentation	5545wave_two_documentation.pdf	4073
Study information and citation	UKDA_Study_5545_Information.htm	33
Unclassified	5545dfe_lsype_confidentiality_agreement_2012.doc	596
READ File	read5545.htm	11

RELATED STUDIES AND GUIDES

Gaining an Overview (2)

- Interactive LSYPE site is a very accessible way to review what variables are available by wave and category

https://www.education.gov.uk/ilsype/workspaces/public/wiki/Welcome

Gaining an Overview (3)

- Example showing wave 1 question to parent about if young person has a health problem or disability and
- Follow up question about whether these problems make it harder to attend school regularly

W1chea1HS - Details

Label : HR: Whether YP has any long-standing illness, disability or infirmity

Dataset : Wave 1 Young Person

Source Question(s) : Question

Universe : {Ask all}

Derivation usages Graph

W1disabYP DV: Whether young person has a disability/long term illness or health problem

Cases, weighted : **15383.8 valid**, 386.2 missing or empty.
Weight used : W1FinWt

Value labels		N(wt)	%(wt)
-1	Don't know	11.5	
1	Yes	2280.1	15 %
2	No	13092.2	85 %

Missing Value labels		N(wt)
-995	Missing history section data - unexplained	83.5
-99	History respondent not interviewed	288.1
-97	History respondent misidentified	14.6

W1chea7HS - Details

Label : HR: Whether these problems make it hard for YP to attend school regularly

Dataset : Wave 1 Young Person

Source Question(s) : Question

Universe : **{If has long-standing illness, disability, infirmity (Chea1=Yes)}**

Derivation usages Graph

W1disabYP DV: Whether young person has a disability/long term illness or health problem

Cases, weighted : **2280.1 valid**, 13489.9 missing or empty.
Weight used : W1FinWt

Value labels		N(wt)	%(wt)
-1	Don't know	11.5	1 %
1	Yes	601.7	26 %
2	No	1667	73 %

Missing Value labels		N(wt)
-995	Missing history section data - unexplained	83.5
-99	History respondent not interviewed	288.1
-97	History respondent misidentified	14.6
-91	Not applicable	13103.6

Data Access (1)

Publically Available Data

• Can access data from waves 1-7 at the UK Data Service

http://ukdataservice.ac.uk/get-data/key-data/cohort-and-longitudinal-studies.aspx

• See the How to access MCS video by John Holmes

https://www.shef.ac.uk/polopoly_fs/1.349153!/file/How_to_access_the_MCS.swf

• Data can also be downloaded from the interactive LSYPE website once you have registered your research project however the process is less smooth

Data Access (2)

Confidential Data

- Confidential data can be accessed through the secure data service (secure lab)

- Necessary if you would like extensive detail on the individual's school, greater detail on their health condition or anything else that might make the individual relatively identifiable

- It takes a little while to register so if you are likely to need data of this sort it is worth identifying whether this is likely to be a necessary step

- Data can only be accessed through the portal rather than downloaded – not all statistical packages are available (MPlus isn't, using R is not easy)

Managing the Data

UK Data Service (Publically available data)

- Download all files into a folder
- Create do file to:
    - open files, keep necessary variables, merge with other data files, clean and create variables

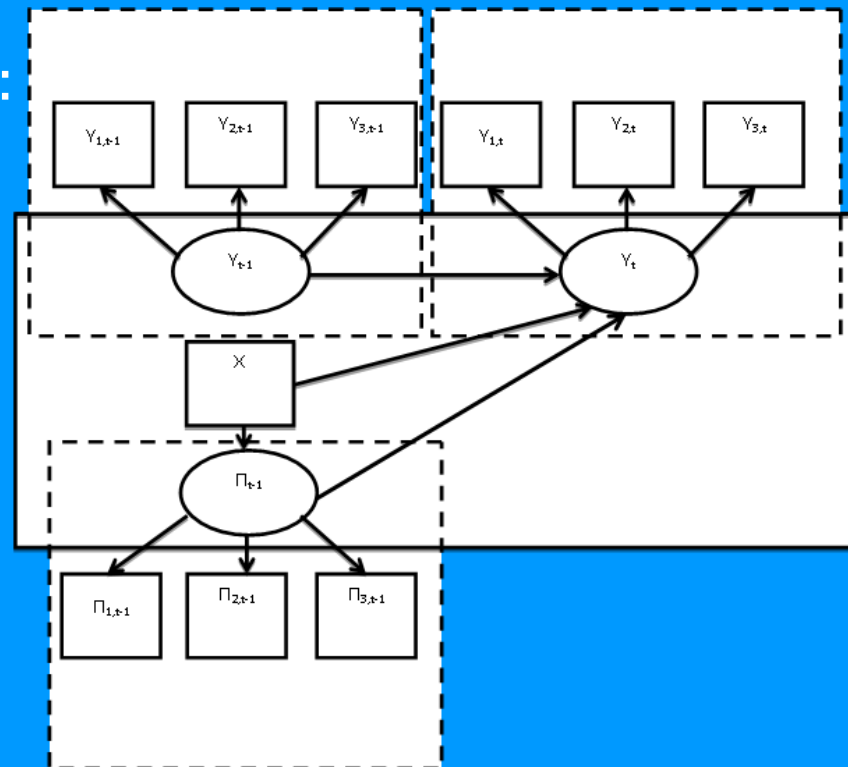**Secure Lab (Confidential data)**

- Your requested Data files are downloaded into a subfolder folder
- You can upload:
    - data files from UK data service
    - Syntax you have used previously to clean and create variables
- You are allocated a sub folder in which to leave your syntax for merging datasets, cleaning variables, running analyses etc.
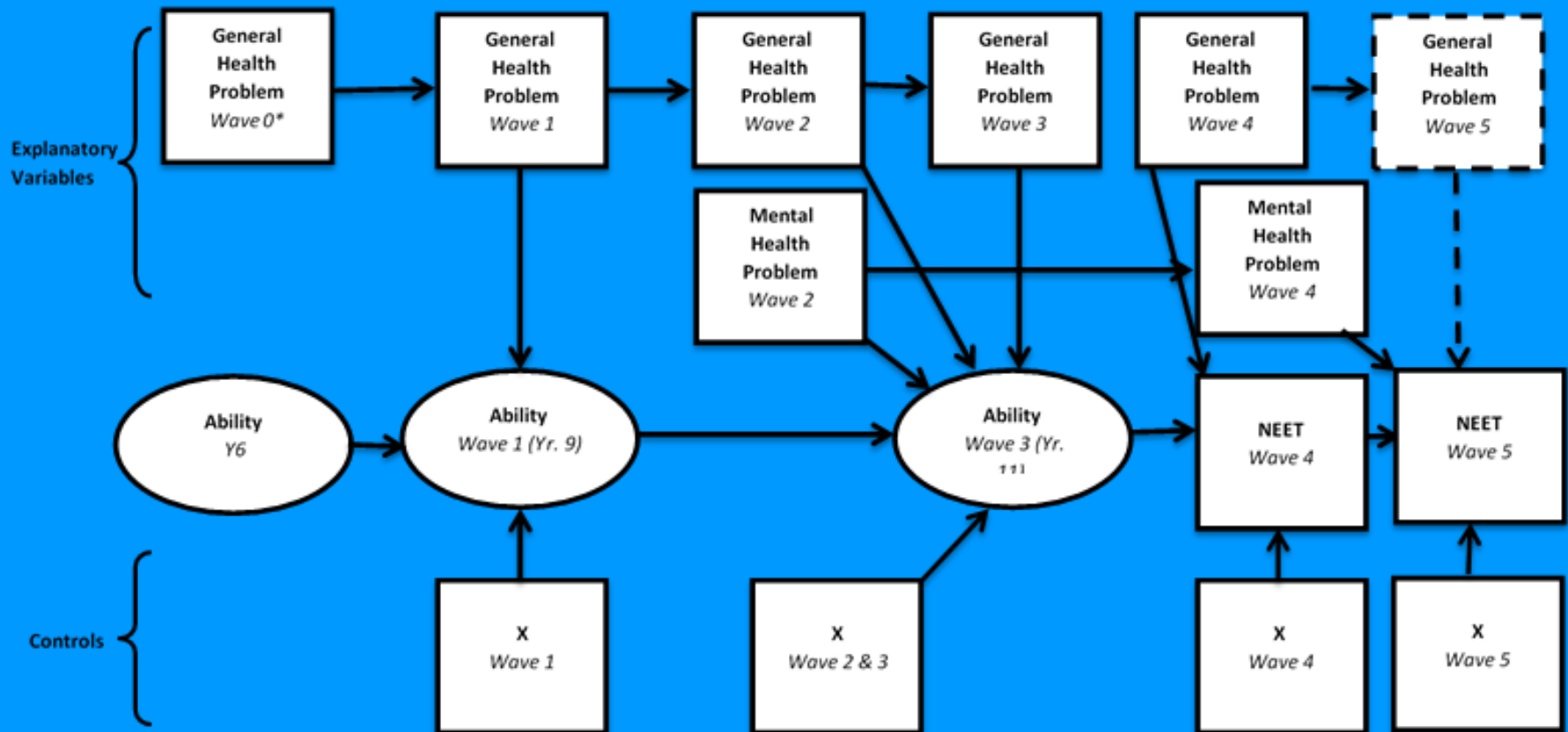
An Example:
The formation of ability in adolescence

- Example uses data from the first 5 waves of the LSYPE (Uses a Structural Equation Modelling Approach)

- General Example of a SEM:

An Example:
The formation of ability in adolescence

Post-compulsory participation and distance (Dickerson and McIntosh, *Urban Studies, 2013)*

- What affects the decision to remain in post-compulsory education is of interest.

- LSYPE provides an extensive list of potential determinants – at the level of the pupil, family, school, local area.

- We considered the distance from the individual's home to their nearest educational institution

Results – by GCSE attainment

LSYPE	Academic study	Vocational study
No GCSEs at grade A*–C		
distance from nearest academic inst.	–0.000 (0.001)	–0.004 (0.010)
distance from nearest vocational inst.	–0.001 (0.001)	–0.003 (0.005)
1–3 GCSEs at grade A*–C		
distance from nearest academic inst.	0.000 (0.005)	0.012 (0.009)
distance from nearest vocational inst.	0.004 (0.002)	–0.010 (0.004)*
4–6 GCSEs at grade A*–C		
distance from nearest academic inst.	–0.038 (0.011)**	0.024 (0.009)**
distance from nearest vocational inst	0.006 (0.004)	–0.008 (0.004)*
7+ GCSEs at grade A*–C		
distance from nearest academic inst.	–0.001 (0.002)	0.002 (0.002)
distance from nearest vocational inst.	0.001 (0.001)	–0.000 (0.001)