

The Longitudinal Study of Young People in England (LSYPE)

Andrew Ledger
& James Halse

Strategic Analysis, DfES

Overview

- The background to LSYPE
- Basic description of the study
- The policy angle

- LSYPE design
- Wave 1 data & illustrative results
- Wave 2 and beyond

Background to LSYPE

- **Gaps in our understanding**
 - Ethnicity, social class, family and parental influence, longer transitions
- **Existing data sources inadequate**
- **Years of development work, consultation & debate over details**
 - Start year, coverage, methodology etc.

Details of LSYPE (1)

- The Department's largest ever research investment
- Longitudinal – will run for up to 10 years
- Interviews at Wave 1 with young people in Year 9 (~30 min)
- Parent(s) also interviewed (45+ min)

Details of LSYPE (2)

- Almost 16 thousand young people and parent(s)
- About 40 thousand interviews
- 650 interviewers
- 3 contractors....
- ...Vast amounts of information

Details of LSYPE (3)

- **Sampled initially from administrative records**
- **Sample boosts for main Minority Ethnic Groups**
- **Also boosts for schools with high prevalence of Free School Meals**
- **Links to admin systems holding exam results**

Details of LSYPE (4)

- **Attitudes of young person and parents**
- **Risk factors (self-completion)**
- **Plans for future – staying on, HE etc.**
- **Background factors**
 - Household SEC
 - Household income & work history
 - Household composition & history
- **Other linked data – local / school level**

Examples of policy angles

- Widening participation
- Social mobility
- The NEET – not in education, employment or training
- Interests of other Departments –eg. DWP interest as our cohort joins the labour market

The sample

- Target population = all pupils in year 9 in England in 2003/04 (with a few exceptions)
- For practical reasons, most of the sample was drawn from year 8 records in 2002/03
- 2 stage sample:
 - Stage 1: sample schools using modified probability proportional to size (PPS) design
 - Stage 2: sample year 9 pupils within selected schools

Sampling schools

- Maintained schools sampled using the Pupil Level Annual Schools Census (PLASC)
- Maintained schools stratified into deprived/non-deprived
- Within each stratum, a size measure was calculated dependent on number of pupils from major ethnic minority groups (Indian, Pakistani, Bangladeshi, Black African, Black Caribbean, Mixed) in year 8 at that school
- A small sample of independent schools also selected

Sampling pupils

- Within each school, selection probabilities were calculated for pupils to ensure issued sample numbers of 1000 from each of the main ethnic minority groups
- Importantly, the way ethnic minorities were boosted means that all pupils within an ethnic group and within a school deprivation stratum have an equal chance of selection

Wave 1 data - response

- **Good overall response rate for the issued sample (74%), but some evidence of non-response bias, e.g.**
 - Black Caribbean R.R. = 67%
 - Indian R.R. = 78%
- **Evidence of non-response bias at school level**

Wave 1 data – weighting (1)

- Weights have been created to account for differential selection probabilities and non-response bias
- The final weights (w1totwt1) range from 0.1 to 9.6
- Use this weight for analysis

Wave 1 Data – weighting (2) young people by ethnic group

	Actual	Weighted
White	10,335	13,399
Mixed	795	418
Indian	1,013	372
Pakistani	940	339
Bangladeshi	722	139
Black Caribbean	576	215
Black African	613	227

Wave 1 data - Design effects

- The two stage design (stratification and clustering), and the differential weights (selection and non-response) lead to quite large design effects

Comparison of confidence intervals around average key stage 3 points scores calculated using usual SPSS method and a robust method (Stata)

	Mean	Strd error	[95% Confidence Interval]	
SPSS usual method	33.95	0.055	33.84	34.06
Robust method	33.95	0.148	33.66	34.24

Wave 1 data – set up

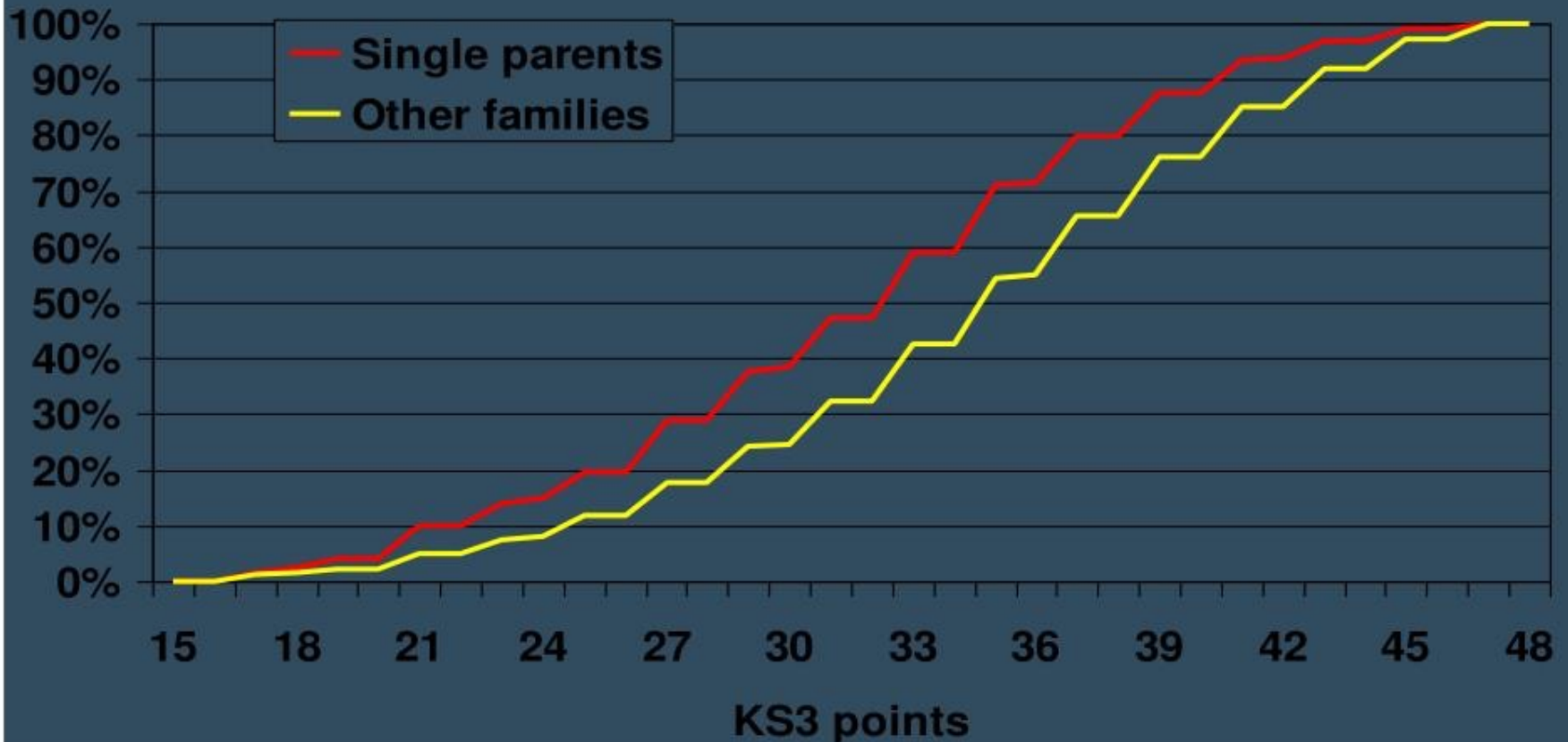
- There are over 1000 variables in the wave 1 survey data alone
- To make the data less unwieldy it's been split into separate files – young person, main parent, second parent and child history – all linkable by survey ID
- Also some hierarchical files, e.g. parental employment histories
- Linked administrative data has not been deposited but can be requested from the Department

Wave 1 data – examples of analysis

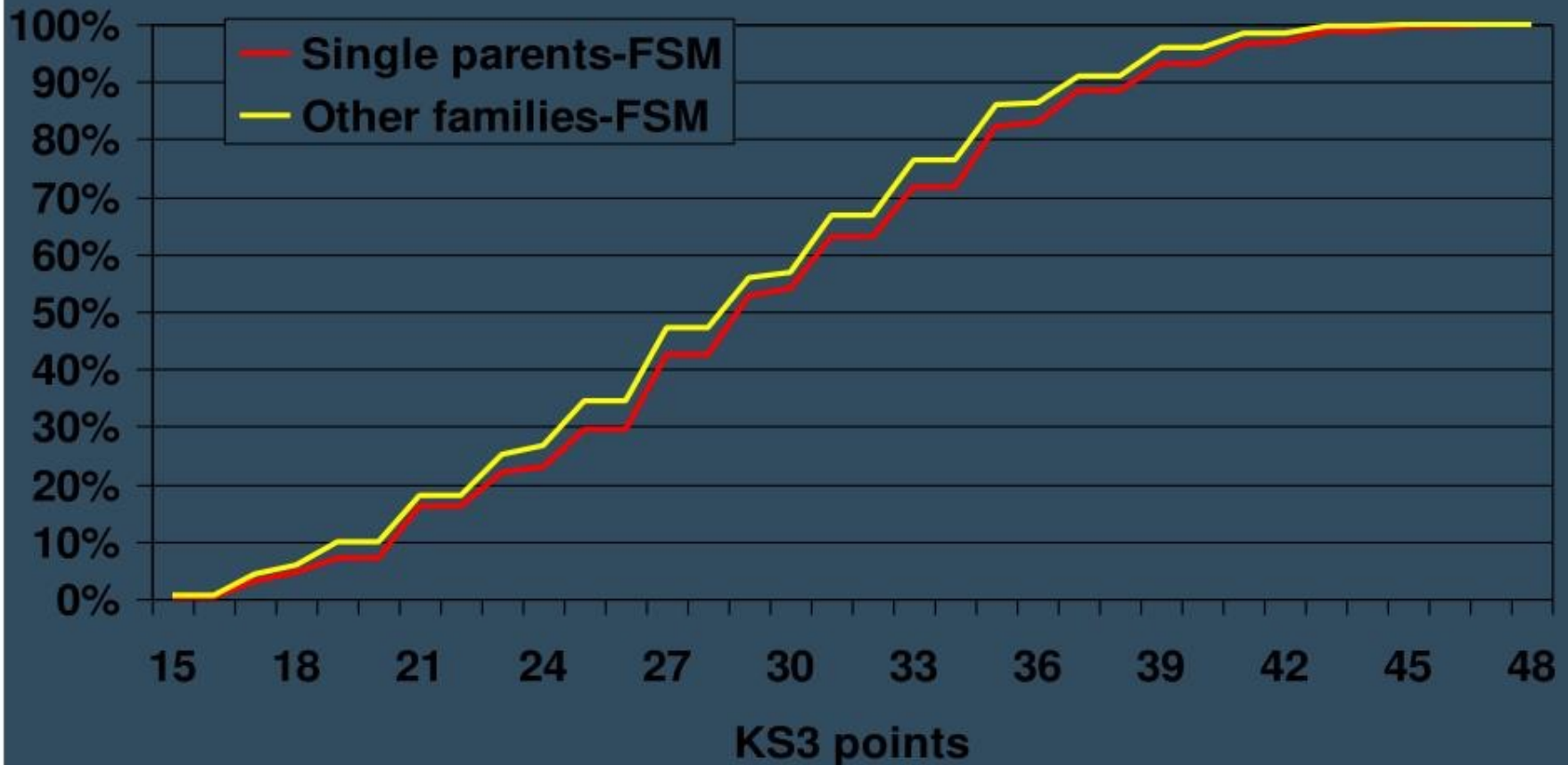
- **Minority ethnic groups (Steve Strand)**
- **The Relationship between Pupil Attainment at Key Stage 3 and Family and Neighbourhood Factors (Andrew Jenkins, Rosalind Levačić, Lars-Erik Malmberg)**
 - Highly significant associations between neighbourhood disadvantage and attainment at Key Stage 3 even after controlling for family background factors
- **Children in lone parent families (DfES internal analysis)**

Wave 1 analysis

Single Parents Look Bad



But Maybe They're Just Poor



Future plans

Data Enhancement project

- **Aim:**
 - Increase the usability and availability of the data
- **Do this by:**
 - Quality assuring and cleaning the data
 - Adding useful derived variables
 - Full documentation
 - Non-response weighting
 - Creating a web-based resource for people to find out more about the data, and to run simple analysis

Data enhancement

- **Updates schedule (provisional):**

- Late summer 2007**

- Wave 1 update
- Wave 2

- Late autumn 2007**

- Wave 3

- Winter 2008**

- Updates to waves 1 to 3
- Starter file

Feedback from users

- What derived variables have you created/would you find useful?
- What should be included in the starter file?
- What strategies have you employed for dealing with missing data?

Questions?

Andrew Ledger:

andrew.ledger@dfes.gsi.gov.uk

0114 259 3639

James Halse:

james.halse@dfes.gsi.gov.uk

0114 259 3868